

# Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption

N. I. Kristiansen,<sup>1</sup> A. Stohl,<sup>1</sup> A. J. Prata,<sup>1</sup> N. Bukowiecki,<sup>2</sup> H. Dacre,<sup>3</sup> S. Eckhardt,<sup>1</sup> S. Henne,<sup>4</sup> M. C. Hort,<sup>5</sup> B. T. Johnson,<sup>5</sup> F. Marengo,<sup>5</sup> B. Neininger,<sup>6</sup> O. Reitebuch,<sup>7</sup> P. Seibert,<sup>8</sup> D. J. Thomson,<sup>5</sup> H. N. Webster,<sup>5</sup> and B. Weinzierl<sup>7</sup>

Received 8 September 2011; revised 25 November 2011; accepted 1 December 2011; published 4 February 2012.

[1] The requirement to forecast volcanic ash concentrations was amplified as a response to the 2010 Eyjafjallajökull eruption when ash safety limits for aviation were introduced in the European area. The ability to provide accurate quantitative forecasts relies to a large extent on the source term which is the emissions of ash as a function of time and height. This study presents source term estimations of the ash emissions from the Eyjafjallajökull eruption derived with an inversion algorithm which constrains modeled ash emissions with satellite observations of volcanic ash. The algorithm is tested with input from two different dispersion models, run on three different meteorological input data sets. The results are robust to which dispersion model and meteorological data are used. Modeled ash concentrations are compared quantitatively to independent measurements from three different research aircraft and one surface measurement station. These comparisons show that the models perform reasonably well in simulating the ash concentrations, and simulations using the source term obtained from the inversion are in overall better agreement with the observations (rank correlation = 0.55, Figure of Merit in Time (FMT) = 25–46%) than simulations using simplified source terms (rank correlation = 0.21, FMT = 20–35%). The vertical structures of the modeled ash clouds mostly agree with lidar observations, and the modeled ash particle size distributions agree reasonably well with observed size distributions. There are occasionally large differences between simulations but the model mean usually outperforms any individual model. The results emphasize the benefits of using an ensemble-based forecast for improved quantification of uncertainties in future ash crises.

**Citation:** Kristiansen, N. I., et al. (2012), Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, 117, D00U11, doi:10.1029/2011JD016844.

## 1. Introduction

[2] An explosive eruption of the Eyjafjallajökull volcano (63.63°N, 19.61°W, 1666 m above sea level (a.s.l.)) on Iceland started on 14 April 2010 and continued for six weeks until the end of May 2010. The eruption was characterized

with a medium volcanic explosivity index (VEI) of 4, of which more than 100 eruptions have been reported since 1500AD [Newhall and Self, 1982]. The erupted total fine ash mass (particles with diameter 2.8–28  $\mu\text{m}$ ) was estimated to be approximately 10 Tg [Schumann et al., 2011; Stohl et al., 2011]. Volcanic ash is a known hazard to aviation and there have been considerable efforts to mitigate the problem using satellite measurements and advanced dispersion models [Prata and Tupper, 2009]. The ash was transported eastward and southwards to the European mainland in the days after the eruption onset and caused closure of airports all over Europe. 100,000 flights were canceled during the eruption period with over 10 million people affected. 8,200 flights were canceled on the first day of the crisis alone [European Commission, 2011].

[3] Prior to the Eyjafjallajökull eruption, European civil aviation authorities (CAAs) imposed a policy of zero tolerance to volcanic ash, meaning that if any “visible” ash was forecasted in the air space, aircraft were re-routed or grounded [International Civil Aviation Organization, 2007].

<sup>1</sup>NILU - Norwegian Institute for Air Research, Kjeller, Norway.

<sup>2</sup>Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, Villigen, Switzerland.

<sup>3</sup>Department of Meteorology, University of Reading, Reading, UK.

<sup>4</sup>Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland.

<sup>5</sup>Met Office, Exeter, UK.

<sup>6</sup>Centre for Aviation, Zurich University of Applied Sciences, Winterthur, Switzerland.

<sup>7</sup>Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany.

<sup>8</sup>Institute of Meteorology, University of Natural Resources and Life Sciences, Vienna, Austria.

Following the initial eruption, the European CAAs divided the air space into regions of high ( $>4000 \mu\text{g}/\text{m}^3$ ), medium ( $2000\text{--}4000 \mu\text{g}/\text{m}^3$ ) and low ( $200\text{--}2000 \mu\text{g}/\text{m}^3$ ) ash contamination [European Commission, 2010, 2011]. Airspace closure is undertaken for high ash contamination, while aircraft are allowed flying with certain restrictions in regions of low and medium levels of ash.

[4] The official forecasts of the transport of the volcanic emissions are provided through the Volcanic Ash Advisory Centers (VAAC) of which the London VAAC has the responsibility for eruptions in the Icelandic region. During the eruption, other national Met Offices and many other research institutes also made forecasts of the ash transport [e.g., Norwegian Meteorological Institute, 2010; Monitoring Atmospheric Composition and Climate, 2010; Barcelona Supercomputing Center, 2010; National Environmental Research Institute, 2010; NILU - Norwegian Institute for Air Research, 2010; Alaska Volcano Observatory, 2010]. The ash concentration limits introduced a new level of accuracy required from the ash transport forecasts; from qualitative ash alerts identifying contaminated regions, to predicting ash concentrations in the air space quantitatively. The main challenge related to this is the lack of knowledge about the source term of the eruption; required as input to the dispersion models. The source term includes parameters like the height of the ash plume, the mass eruption rate, the duration of the eruption and mass fraction of fine ash (small particles which can remain in the cloud for many hours or days and thus be transported far from the source) [Mastin et al., 2009]. Also the distribution of the ash in the eruption column is an important source parameter. Because of the atmosphere's thermal structure and wind profiles, the ash is likely to be emitted at certain heights from which it laterally spreads out. Then, the ash is transported in different directions due to wind shear and depending on the time and height of the emissions. This yields significant differences in the 3-dimensional spatial distribution, and hence uncertainties in the predicted areas where ash concentration limits will be exceeded.

[5] There exist various methods for estimating the source term of a volcanic eruption. The most common approach is based on observations of plume heights, e.g., from weather radar, which are fed into an empirical relationship linking the total mass emission rate to the eruption column height [Sparks et al., 1997; Mastin et al., 2009]. This relationship has large uncertainties and a wide range of total mass emitted can fit one plume height. Also the fraction of fine ash, used for model simulations, varies a lot between volcanoes and eruptions. Furthermore, the heights at which ash is detrained from the vertical column above the volcano are not given by this method. The common set-up, currently used at the London VAAC, assumes a uniform eruption column where ash is released uniformly from the volcano vent up to the reported plume heights [Witham et al., 2007; Webster et al., 2012]. However, this is a crude simplification of the source term and, if the true distribution can be assessed, this should enable more accurate predictions.

[6] In addition to assumptions about the source term there are a number of other factors yielding uncertainties in the model predictions of volcanic ash transport. These include uncertainties in the dispersion model itself, the meteorological input data used for driving the model

and assumptions made for the particle size distribution of the ash.

[7] In this paper we present source term estimates for the ash emissions from the Eyjafjallajökull eruption using an inversion algorithm which is based on modeled scenarios of the ash emissions constrained by satellite column data and additionally constrained by a priori information. Similar results are also presented by Stohl et al. [2011] for a subset of these analyses. Here we present results based on different models, run on different meteorological data, and evaluate the model simulations with a large set of independent ash concentration measurements both from surface and aircraft instruments. The results are valuable for assessing the dispersion models' capability to accurately predict ash concentrations far from the ash emission source and for evaluating some of the uncertainty factors related to ash forecasting.

## 2. Methods

[8] We have used two different models, and several different model set-ups to simulate the transport of the ash emissions from the Eyjafjallajökull eruption. The emission scenarios (source-receptor sensitivities) obtained by these simulations are used as input to an inversion algorithm which incorporates the modeled ash emissions and satellite observations of the ash cloud to give an optimal estimate of the ash release from the volcano as a function of time and height. This section presents first the different models used and the set-up of the transport simulations, followed by a short description of the satellite data used and the inversion method of estimating the source term. Finally, an overview over the measurement data used for model validation is given.

### 2.1. Transport Models and Simulations

[9] Two different Lagrangian particle dispersion models, FLEXPART and NAME, are used for simulating the ash transport for the Eyjafjallajökull eruption. These models calculate the dispersion by tracking model particles through the modeled atmosphere. The particles move with the resolved wind described by the meteorology input and by parameterized small-scale motions and processes like gravitational settling which are not resolved by the meteorology data input. The modeled ash concentrations are calculated on a prescribed grid averaged in space and time. FLEXPART and NAME are two models of the same "type" (i.e., Lagrangian and not Eulerian models) with the main difference being how the models are initialized and also the parameterizations of the removal processes considered (e.g., wet deposition), the convection schemes, and turbulence and other parameterizations.

[10] Aggregation is caused by collision of ash particles and their ability to adhere, and can result in efficient removal of ash from the atmosphere due to the larger sedimentation velocity of the aggregates. This process is particularly efficient in the case of "wet" eruptions and whenever ice forms. Some attempts at modeling the aggregation process have been made by solving the stochastic coagulation equation [e.g., Costa et al., 2010; Folch et al., 2010] but these formulations are too slow for inclusion into operational dispersion models. Neither FLEXPART nor NAME considers aggregation processes in the transport simulations.

### 2.1.1. FLEXPART

[11] The FLEXPART model was first described and validated by *Stohl et al.* [1998] with data from continental scale tracer experiments and is now used for a large range of applications, including simulating the dispersion of volcanic plumes [*Prata et al.*, 2007; *Eckhardt et al.*, 2008; *Kristiansen et al.*, 2010; *Stohl et al.*, 2011].

[12] The FLEXPART simulations for the Eyjafjallajökull eruption were driven with 3-hourly meteorological data from two different centers; the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). The ECMWF analyses data have  $0.18^\circ \times 0.18^\circ$  horizontal resolution and 91 vertical model levels, while the GFS analyses data have  $0.5^\circ \times 0.5^\circ$  resolution with 26 pressure levels. The simulations take into account wet and dry deposition [*Stohl et al.*, 2005] and particle gravitational settling [*Näslund and Thaning*, 1991]. Further details on the FLEXPART simulations are given by *Stohl et al.* [2011].

### 2.1.2. NAME

[13] The Numerical Atmospheric-dispersion Modeling Environment (NAME) model [*Jones et al.*, 2007] was developed by the UK's Met Office following the Chernobyl nuclear accident and is now used for a wide range of applications, including volcanic ash modeling [*Witham et al.*, 2007; *Webster et al.*, 2012; *Dacre et al.*, 2011; *Leadbetter and Hort*, 2011; *Devenish et al.*, 2012]. NAME is the dispersion model of the London VAAC and decisions on air space closure during the Eyjafjallajökull eruption were based on NAME predictions.

[14] NAME was driven by meteorological data from the global version of the Met Office's weather forecast model MetUM (the Met Office Unified Model) with a temporal resolution of 3 h, a horizontal resolution of  $0.35^\circ \times 0.23^\circ$ , and 59 vertical levels up to 29 km altitude. The NAME runs are post event reruns using analyzed meteorological data (analysis every 6 h alternated with 3 h forecasts). NAME also considers wet and dry deposition and gravitational settling of particles [*Maryon et al.*, 1999; *Webster and Thomson*, 2012].

[15] Gravitational settling is calculated in both FLEXPART and NAME assuming spherical particles and based on the Stoke's law for fall velocity for small particles, including a Cunningham slip coefficient which increases the fall rate of small particles at high elevation. For larger particles (Reynolds number  $>1$ ) the sedimentation velocity is calculated from the particle density and diameter using the Reynolds number dependent drag coefficient.

### 2.1.3. Model Runs of Ash Emissions

[16] We performed thousands of model simulations with different scenarios of the ash emissions to produce the input to the inversion algorithm. Each model run simulated the emissions from one of 19 height levels of the eruption column, and one of 328 three-hourly time steps during the whole 41-days eruption period. Every simulation was run forward in time for six days and produced hourly averaged total atmospheric ash columns with  $0.25^\circ \times 0.25^\circ$  horizontal resolution in the domain  $30^\circ\text{W}$  to  $30^\circ\text{E}$  and  $40^\circ\text{N}$  to  $70^\circ\text{N}$ . Each simulation carried one unit mass of ash spread over thousands of particles. Thus, these results give the sensitivity of the total ash column loadings to the ash source in one

single emission grid box. More details on these emission scenario simulations are given by *Stohl et al.* [2011].

[17] The model particles were distributed over a particle size distribution which was determined by fitting a lognormal distribution to a ground sample of the ash taken 60 km from the volcano, and guided also by qualitative comparisons with an airborne sample from the DLR Falcon research aircraft flying at a 450 km distance from the volcano [see *Stohl et al.*, 2011, Figure 1]. One important consideration is that satellite observations are limited to observe particles within a certain size range ( $2\text{--}32\text{ }\mu\text{m}$  diameter). Since the inversion algorithm compares model values and satellite observations, only this part of the particle size range was used for the model simulations. Both FLEXPART and NAME used the same particle size distribution.

### 2.1.4. Long-Range Transport Simulations

[18] The estimated source terms from the inversion were used to perform long-range transport simulations of the ash emissions using both the FLEXPART and NAME models. The emissions were released as a non-uniform line source above the volcano (i.e., no horizontal extent). The output from both models was on a  $0.25^\circ \times 0.25^\circ$  horizontal grid with 250 m vertical resolution and with hourly averaged ash concentrations (with an integration time of 5 min). The particle size distribution used for these simulations was extended to a larger particle size range ( $0.25\text{--}250\text{ }\mu\text{m}$  diameter) [see *Stohl et al.*, 2011, Figure 1]. The distribution has a primary mode at  $10\text{ }\mu\text{m}$  diameter and a secondary mode at  $180\text{ }\mu\text{m}$ . Ash particles with diameter less than  $20\text{ }\mu\text{m}$  are fine enough to remain in the air for a longer time, and thus can survive transport to Europe [*Schumann et al.*, 2011]. It is noted however that aggregation can affect the true particle size distribution in the distal ash clouds.

[19] In addition to the long-range transport simulations using the source term from the inversion, a simpler source term was used for one FLEXPART simulation run on ECMWF meteorological data. This simulation has the same output definitions as described above, but the emissions of ash were uniformly distributed in the vertical above the volcano. Also the emission rate was based on the relationship between observed plume heights and mass emission rate as described by *Mastin et al.* [2009]. The 3-hourly averaged observed plume heights were taken from radar observation at the Keflavik airport [*Arason et al.*, 2011]. In addition, an assumption of 10% fine ash fraction was used. This yields a total release of 34.4 Tg of fine ash for the whole eruption period used in this simulation. This model simulation is referred to as the uniform simulation and is similar to what is used operationally at the London VAAC.

[20] FLEXPART stores the model output separately for each particle size, which allows for comparisons of modeled size distribution to measured size distributions. The current version of NAME does not have this capability directly. However, size distributions from NAME were retrieved from other simulations by *Dacre et al.* [2011], who ran the model several times with different size bins. These simulations use a different source term and initial size distribution than the NAME simulations presented in this paper. Specifically, the initial size distribution is the default one used by NAME for volcanic ash and is given by *Dacre et al.* [2011, Table 1], and the simulations use a uniform vertical source profile. Despite the different set-up, the size distributions retrieved from

these NAME simulations are thought to be valuable for comparison to both the measured size distributions and the FLEXPART-derived size distributions. In comparing the modeled and measured size distribution we are primarily looking at the relative shapes of the size distributions rather than the absolute values.

## 2.2. Satellite Data

[21] Measurement data from the geosynchronous Meteosat Second Generation (MSG) Spin-stabilized Enhanced Visible and Infrared Imager (SEVIRI) was used in the inversion algorithm to constrain the modeled volcanic ash emissions. Measurements taken in infrared channels were used to retrieve total atmospheric column ash mass loadings ( $\text{g/m}^2$ ) with an estimated error of 40–60% [Wen and Rose, 1994]. The satellite retrievals are only sensitive to ash with particle diameters from 2 to 32  $\mu\text{m}$ . For the inverse modeling, the 15 min pixel-by-pixel mass loading retrievals were time-averaged to hourly time-intervals and re-gridded to a  $0.25^\circ \times 0.25^\circ$  grid to provide mass loading values over the same domain as the modeled ash columns ( $30^\circ\text{W}$  to  $30^\circ\text{E}$  and  $40^\circ\text{N}$  to  $70^\circ\text{N}$ ). Detailed information about the retrievals is given by Stohl *et al.* [2011] and A. J. Prata and A. T. Prata (Eyjafjallajökull volcanic ash concentrations determined from SEVIRI measurements, submitted to *Journal of Geophysical Research*, 2011).

## 2.3. Source Term Estimations

[22] The inversion method used for the ash source term estimations is based on work of Seibert [2000] and has been further developed and tested with different volcanic eruptions [Eckhardt *et al.*, 2008; Kristiansen *et al.*, 2010; Stohl *et al.*, 2011] and for determining greenhouse gas emissions [Stohl *et al.*, 2009]. All details about the algorithm are described in these papers and here we only give a short summary of the method and the set-up.

[23] The inversion algorithm combines the modeled emission sensitivities and satellite measurement data as described in the previous sections. To render the solution of the inversion more stable, a priori emissions need to be used. As in the work of Stohl *et al.* [2011] we used the 1-D model for convective volcanic plumes, PLUMERIA [Mastin, 2007], which uses actual atmospheric conditions taken from ECMWF data and observed plume heights [Arason *et al.*, 2011; Jakobsdóttir *et al.*, 2010; Volcanic Ash Advisory Centre London, 2010] to determine the a priori emissions as a function of height and time. Furthermore, we assumed that 10% of the erupted mass was fine ash in the size range to which the satellite measurements are sensitive. Note that these a priori emissions are a more sophisticated representation of the source term than that used for the operational set-up at the London VAAC. The London VAAC use the empirical relationship between observed plume heights and the eruptive mass from Mastin *et al.* [2009], with a uniform vertical ash distribution. We note however that the use of a more diffuse profile, as used by the London VAAC, may have advantages in a risk assessment context where one may not want to concentrate the release over a narrow range of heights unless one is confident of the choice of height range.

[24] The total columns of ash from each of the 6232 modeled ash emission scenarios (3 hourly releases over 19 height levels and for 41 days) are scaled with the a priori

emissions, and subsequently for each hourly time step every satellite data pixel is compared with 912 model values (from 3 hourly releases up to 6 days back in time for 19 height levels). Approximately 2 million satellite observations were used for the whole inversion, which yields around 1.8 billion satellite-model comparisons. We performed three different inversions with model input from (1) FLEXPART run on ECMWF meteorological data, (2) FLEXPART run on GFS data and (3) NAME run on MetUM data. The results of the inversions are a posteriori source terms for the optimized ash emissions.

[25] The inversion also considers the uncertainties in the various inputs; a priori emission uncertainties, and errors in the observations and the model. These are important parameters as they are used to constrain the result by allowing the inversion to substantially change the a priori emissions, while still being guided by the a priori estimate. Furthermore, if the observations do not provide enough constraints on the emissions, the solution will remain close to the a priori estimate. The uncertainties applied in the inversion set-up are identical to those used by Stohl *et al.* [2011]. The uncertainties are further assessed by Seibert *et al.* [2011] where uncertainties for the a posteriori emissions are introduced.

[26] The set-up of the inversions used in this study differs slightly from the inversion set-up and results presented by Stohl *et al.* [2011]. Here, only the SEVIRI satellite data were used as only these data would be available on a near-real time basis. For the FLEXPART inversions presented by Stohl *et al.* [2011], IASI (MetOp Infrared Atmospheric Sounding Interferometer) satellite data were also used, which were tuned to the SEVIRI data, and the average over two inversions, from the ECMWF- and GFS-based model data was presented. The results were not significantly different from the results presented here using only SEVIRI data. Stohl *et al.* [2011] also show that the inversion is not very sensitive to the model meteorological data input (ECMWF or GFS) or the satellite data used (SEVIRI or IASI or both together). However, the sensitivity to using a different dispersion model was not considered, but is presented here. Both the NAME and FLEXPART inversions used the gridded absolute model uncertainties as described by Stohl *et al.* [2011] which are estimated by the differences in modeled ash columns from the ECMWF-based and GFS-based simulations.

## 2.4. Measurement Data

[27] Measurement data of volcanic ash from three research aircraft and one surface measurement station were used to evaluate the inversion results and the long-range transport simulations.

[28] The United Kingdom's Facility for Airborne Atmospheric Measurements (FAAM, <http://www.faam.ac.uk/>) BAe-146 research aircraft was deployed for twelve flights over the UK and the surrounding seas between 20 April and 18 May 2010, to measure the volcanic ash from Eyjafjallajökull [Marenco *et al.*, 2011; Turnbull *et al.*, 2012; B. Johnson *et al.*, In-situ observations of volcanic ash clouds from the FAAM aircraft during the eruption of Eyjafjallajökull in 2010, submitted to *Journal of Geophysical Research*, 2011]. The aircraft was equipped with various optical particle counters including the Cloud and Aerosol Spectrometer (CAS) and the Passive Cavity Aerosol Spectrometer Probe (PCASP) counting and sizing particles in the

size range 0.1–50  $\mu\text{m}$ . The in situ spectrometer instruments count and size the particles based on the detection of the amount of light scattered by the single particles. Thus, the instruments did not measure the mass concentrations of the ash particles directly. The particle number size distributions from CAS and PCASP were used to estimate the total aerosol mass concentration with an assumed ash mode density of 2.3  $\text{g}/\text{cm}^3$  and with a refractive index based on mineral dust. Only the CAS mass (covering the size range 0.6–50  $\mu\text{m}$ ) was used as an estimate for the ash mass, since the PCASP mass was thought to be dominated by secondary aerosols (e.g., sulfate). The ash mass concentrations have an uncertainty of a factor of two. All ash mass derivation methods for the BAe-146 data are described by Johnson et al. (submitted manuscript, 2011). The BAe-146-measured size distributions were subject to an additional level of screening where data points with CAS aerosol mass  $<20 \mu\text{g}/\text{m}^3$  were rejected to avoid background aerosol from biasing the size distribution since some flights contained substantial ash-free sections. In this study we compare the measurements to the modeled along-flight mass concentrations and particle size distributions.

[29] The BAe-146 was also equipped with an on-board light detection and ranging (lidar) instrument operating in the near ultraviolet (355 nm). The lidar-derived aerosol extinction was converted to ash concentrations by combining the lidar measurements with the size-distribution information derived from the in situ probes [Marenco et al., 2011]. The converted ash mass concentrations have an estimated uncertainty of a factor of two, and the vertical resolution of the data was 45 m with an integration time of 1 min, which corresponds to a  $\sim 9$  km footprint as the aircraft traveled at a speed of  $\sim 150$  m/s. Only data beyond 300 m below the aircraft were considered (full receiver-emitter overlap) and the data were screened for clouds in that every time a cloud top was found, the cloud and everything beyond it was removed from the data set. Here, the retrieved ash concentrations from the aircraft's lidar are directly compared with the ash concentrations predicted by the models.

[30] The Falcon research aircraft of the Deutsches Zentrum für Luft- und Raumfahrt (DLR) performed measurements in the volcanic ash clouds from Eyjafjallajökull during 17 flights between 19 April and 18 May 2010 [Schumann et al., 2011; Turnbull et al., 2012]. The instruments onboard the Falcon aircraft were similar to those of BAe-146 and included a Doppler wind lidar (2  $\mu\text{m}$  wavelength) and in situ instruments for measuring aerosol microphysics properties, chemical species and meteorological parameters. The optical laser aerosol Forward Scattering Spectrometer Probe (FSSP-300) counts and sizes the particles in the size range 0.1–30  $\mu\text{m}$  (coarse mode particles) depending on the refractive index, which are considered to represent the ash mass. The measured particle number size distributions were used to derive the mass concentrations for a given particle refractive index (case M presented by Schumann et al. [2011] is used) and ash particle density (2.6  $\text{g}/\text{cm}^3$ ). To provide a uniform analysis Schumann et al. [2011] assumed the same refractive index for all plume penetrations. Furthermore, the refractive index was kept constant over the whole size range. However, Schumann et al. [2011] pointed out that they expect a size dependence of the refractive index, with larger particles ( $>1 \mu\text{m}$ ) being less absorbing than smaller particles.

Therefore, they assumed that the true values may be between the results for cases L and M presented in their study. The derived mass concentration results have an uncertainty of a factor of two. For this study we compare measured and modeled along-flight mass concentrations and particle size distributions.

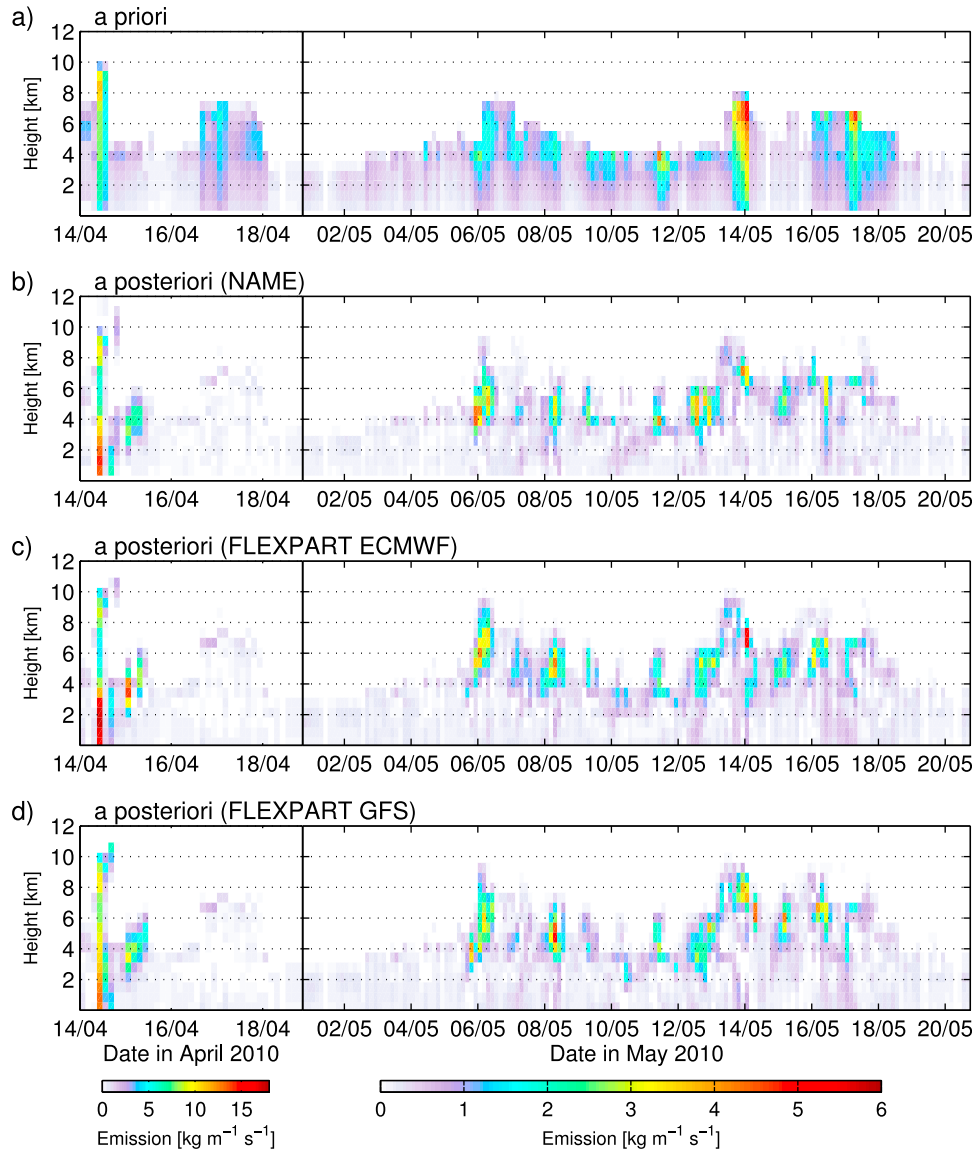
[31] In addition, we have used the vertical profiles of attenuated aerosol backscatter from the lidar onboard the Falcon aircraft to evaluate the height of the ash clouds. The lidar measurements were obtained from an altitude of 400 m below the aircraft to the ground with a vertical resolution of 100 m. The horizontal resolution of the backscatter signal profiles was 150–200 m. The backscatter signal was range-corrected and depends on the vertical profile of the atmospheric backscatter and extinction coefficient, which both depend on the particle (cloud, ash and other aerosol) content of the atmosphere, their size distributions and scattering properties. A mass-conversion of the lidar data was not available yet, thus only qualitative comparisons to the modeled ash clouds' positions are presented here.

[32] A third research aircraft, the Swiss DIMO aircraft (Diamond Aircraft HK36 TTC-ECO) conducted six measurement flights over the Alps in April and May 2010, equipped with two optical particle counters (OPC) for aerosol number measurements. During the flights in April, the pre-installed MetOne OPC (Model 4903, Hach Analytics Inc., USA) was applied, measuring in the size range 0.3–0.5  $\mu\text{m}$  and  $>0.5 \mu\text{m}$ . From 18 May onwards, a Grimm OPC (Model 1.108, Grimm GmbH, Germany) was deployed, with a better size resolution in the range 0.3–10  $\mu\text{m}$  and an optimized sampling line. With these improved conditions, the measurements allowed for aerosol mass concentration estimates using an assumed ash mode density of 2.65  $\text{g}/\text{cm}^3$  and an average complex refractive index for volcanic ash of  $1.54 + 0.005i$  (at the OPC laser wavelength of 780 nm) [Bukowiecki et al., 2011]. The uncertainty for these mass concentration estimates reached up to 50–60%. The retrieved ash mass concentrations and size distribution for 18 May were used for model-comparisons.

[33] The surface measurement station Jungfraujoch in Switzerland (46.55°N, 7.99°E, 3580 m a.s.l.) observed increased particulate matter (PM10) values during two episodes in April and May 2010. The measurements are described in detail by Bukowiecki et al. [2011]. Particle size distributions for larger particles were measured with an OPC that covers a measurable particle size range of 0.3–15  $\mu\text{m}$ . The uncertainties in the mass concentrations estimates are less than 10%. We have utilized these measurements for comparison to the modeled ash concentrations and particle size distributions.

[34] Finally, we also combine all in situ measured ash concentrations from all the twelve BAe-146 flights, seventeen Falcon flights, one DIMO flight and the observations from Jungfraujoch to perform a statistical comparison of the modeled and measured ash concentrations.

[35] When comparing the models with the aircraft in situ measurements, the modeled ash concentrations are extracted from the 3-D space grid with  $0.25^\circ \times 0.25^\circ$  horizontal resolution and the 1-hourly averaged model output. The in situ measurements typically have a sampling time of 1–10 s and are taken over a smaller spatial extent than one single 3-D grid box of the model output grid. Thus, the modeled values



**Figure 1.** Volcanic ash emissions from Eyjafjallajökull volcano as a function of altitude (above the volcano vent, 1666 m a.s.l) and time for the period 14–18 April and 1–20 May 2010. (a) A priori emissions, (b) a posteriori emissions from the inversion using the NAME model input run on MetUM meteorology, (c) a posteriori emissions from the inversion using the FLEXPART model input run on ECMWF meteorology and (d) a posteriori emissions from the inversion using the FLEXPART model input run on GFS meteorology. The a posteriori emissions are constrained using SEVIRI satellite data.

do not represent such instant point concentrations as are measured. For the model-measurement comparisons, the set of in situ measured and modeled values is then averaged over five minutes.

[36] Notice that there are differences in the assumed particle densities for the measurement retrievals and model simulations. While the FLEXPART and NAME model simulations assumed  $3.0 \text{ g/cm}^3$  and  $2.75 \text{ g/cm}^3$ , respectively, the BAe-146 retrievals used  $2.3 \text{ g/cm}^3$ , Falcon  $2.6 \text{ g/cm}^3$ , DIMO and the Jungfraujoch data  $2.65 \text{ g/cm}^3$ . The PLUMERIA calculations used  $2.5 \text{ g/cm}^3$  and the SEVIRI satellite retrieval  $2.6 \text{ g/cm}^3$ . It can also be noted that the operational NAME runs at the London VAAC assume  $2.3 \text{ g/cm}^3$ , while in the work of *Mastin et al.* [2009] a typical

value is  $2.5 \text{ g/cm}^3$ . It is reported that ash particle density for different volcanic eruptions, periods, and particles can vary from  $0.7$  to  $3.2 \text{ g/cm}^3$  (<http://volcanoes.usgs.gov/ash/properties.html#density>). The different particle densities used in this study are thought not to be significant for the overall results.

### 3. Results

#### 3.1. Source Term

[37] The ash emissions for the Eyjafjallajökull eruption as a function of height and time are presented in Figure 1; the a priori emissions derived with the PLUMERIA model (Figure 1a), and the a posteriori emissions constrained



**Table 1.** Total Inverted Mass (Tg) for the Source Terms in Figure 1 (for Particles in the Size Range 2.8–28  $\mu\text{m}$  Diameter) and the Exact Total Ash Mass That Is Emitted in the Model Simulations Using the Source Terms When the Particle Size Distribution Is Extrapolated to a Wider Size Range (Particle Diameter 0.25–250  $\mu\text{m}$ )

	Model (Met Data)				
	All Models	NAME (MetUM)	FLEXPART (ECMWF)	FLEXPART (GFS)	FLEXPART (ECMWF)
Simulation	a priori	a posteriori	a posteriori	a posteriori	uniform
Total inverted mass	$11.5 \pm 11.9$	$6.8 \pm 4.2$	$7.9 \pm 5.2$	$8.1 \pm 5.1$	-
Total mass emitted in model simulations	19.1	11.0	13.1	13.4	34.4

by satellite data for the inversions using NAME with MetUM-meteorology (Figure 1b), FLEXPART with ECMWF-meteorology (Figure 1c) and FLEXPART with GFS-meteorology (Figure 1d) are shown. The source terms have a time resolution of three hours and a vertical resolution of 650 m. The period 19 April to 1 May is not shown as there was very little ash emitted in this period [Stohl *et al.*, 2011].

[38] Generally through the eruption period, the a posteriori ash emissions are reduced compared to the a priori emissions, and slightly shifted to higher emission altitudes. Emissions below about 2 km altitude above the volcano vent are much reduced compared to the a priori. The same temporal and vertical behavior of the a posteriori emissions was also found by Stohl *et al.* [2011].

[39] The most prominent difference between the a priori and a posteriori emissions are found around 17 April with a large reduction of the a priori ash emissions. The difference may be due to collapsing eruption plumes as this was a period of the eruption when the eruption plumes were collapsing more strongly than during other periods (as visible from e.g., web cameras). This means that more of the ash fell out more quickly, and was thus not transported as far downwind and observed by satellites.

[40] The May episode shows large a posteriori emissions between 4 and 7 km altitude above the volcano vent, especially on 6, 8 and 12–14 May. The emissions are also confined to a smaller vertical extent than in the a priori, and most ash is emitted near the top of the eruption column.

[41] The three different inversions using model input from NAME with MetUM-meteorology, FLEXPART with ECMWF-meteorology and FLEXPART with GFS-meteorology, show very similar a posteriori results. All inversions give a clear reduction of the emissions around 17 April, also the emission pulses during the May events are strongly correlated. A noticeable difference occurs on 6 May when the FLEXPART-ECMWF based inversion puts the emissions to a slightly higher altitude than the NAME-MetUM and FLEXPART-GFS based inversions. Also on 12–14 May there are small differences between the a posteriori emissions.

[42] Several sensitivity experiments were conducted by Stohl *et al.* [2011], for comparing the FLEXPART ECMWF-based and GFS-based inversions, and also for the sensitivity of the satellite data input. The results showed rather robust a posteriori emissions. The results presented here demonstrate that the results are robust to using data from different dispersion models and meteorological models.

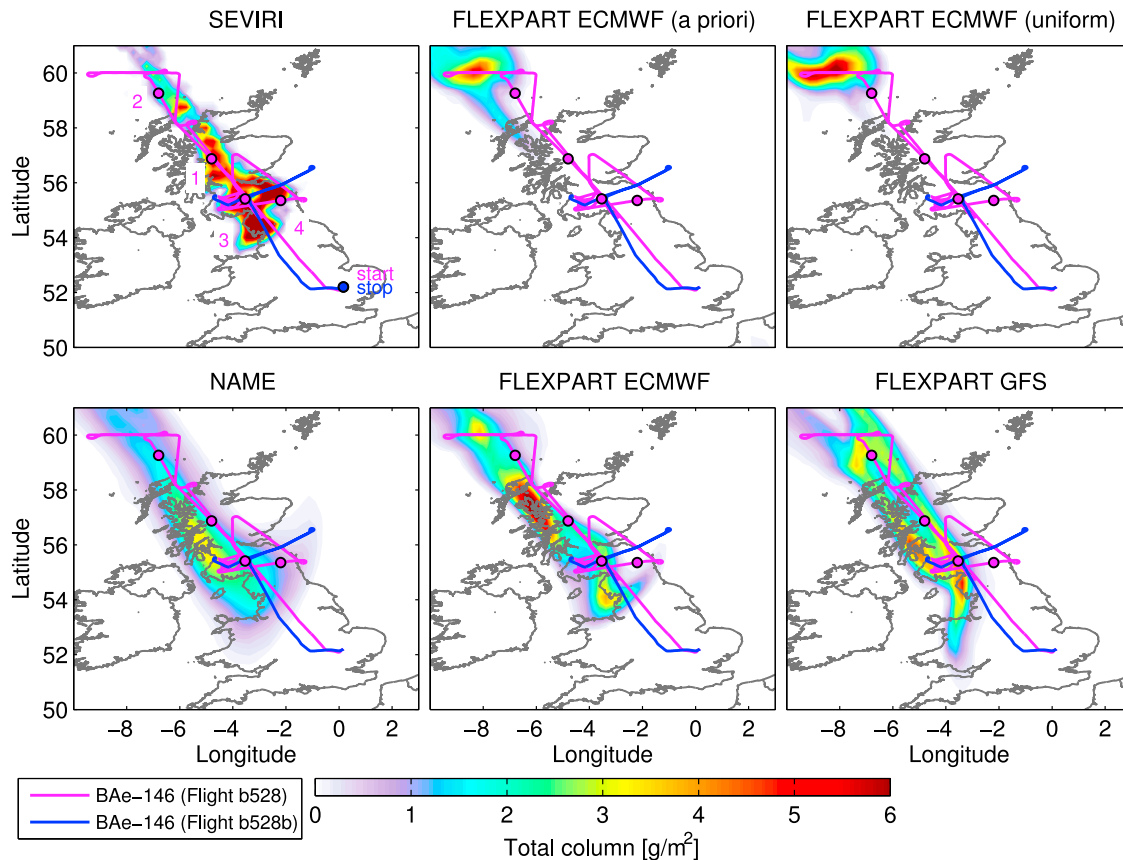
[43] The total fine ash mass of the a priori and the a posteriori source terms are summarized in Table 1. The a posteriori source terms are much reduced compared to the a priori estimate. Also the uncertainties are reduced. The

uncertainty estimations are described by Stohl *et al.* [2011] and Seibert *et al.* [2011]. The two different inversions using FLEXPART model output show very similar total mass results ( $\sim 8.0 \pm 5.1$  Tg) as presented by Stohl *et al.* [2011] ( $8.3 \pm 4.2$  Tg). The NAME inversion gives lower total mass ( $6.8 \pm 4.2$  Tg) than the two other inversions based on FLEXPART model data. The reason for this is unclear but it is likely related to differences in the two models for the removal of particles in the atmosphere (e.g., deposition and gravitational settling). The mean over the three a posteriori source terms is  $7.6 \pm 4.8$  Tg. For the long-range transport simulations using the a priori and a posteriori source terms, the initial size distribution used for the models are extended to a wider range of particles sizes (0.25–250  $\mu\text{m}$  diameter) which yields higher emissions of ash than the inverted mass. Table 1 also gives the ash mass emitted in each model simulation. The total mass emitted in reality by the volcano was even higher and included particles larger than 250  $\mu\text{m}$  (also aggregates) that were quickly removed by sedimentation.

### 3.2. Comparisons of Measured and Modeled Ash Concentrations

[44] The ash concentrations as simulated by FLEXPART and NAME are compared with independent observations from both surface and aircraft measurements. There are seven model simulations in total; three simulations initiated with the a priori emissions of Figure 1 for NAME run on MetUM data, FLEXPART run on ECMWF data and FLEXPART on GFS data; three simulations with the a posteriori emissions for the same models and meteorological input data; one simulation based on a simple source term assuming uniform distribution of the ash within the eruption column (as described in section 2.1.4). The three a posteriori simulations are shown and discussed separately to evaluate how well the models perform in simulating the ash dispersion. Also, for each model grid cell the mean ash concentrations over the three a priori model simulations are calculated, as well as the mean over the three a posteriori simulations, and used as key comparison to the measurement, referred to as the mean modeled a priori or a posteriori ash concentrations.

[45] Three case studies are presented with comparisons of the modeled ash concentrations and measurements from different aircraft campaigns. On 14 May the BAe-146 research aircraft sampled the ash cloud located over Northern England. The ash cloud was further transported across the North Sea where it was sampled by both the BAe-146 and the Falcon on 17 May. On 18 May three different aircraft (BAe-146, Falcon and DIMO) sampled the ash cloud that stretched from the North Sea down to the Alps. Along-track ash concentrations, lidar observations and particle size distributions are presented for each model-measurement



**Figure 2.** Observed and modeled total ash columns on 14 May 2010 between 11:00–12:00 UTC. The flight track for the two BAe-146 flights on 14 May (flight b528 from 10:06–15:41 UTC, flight b528b from 17:28–19:17 UTC) are shown as magenta and blue lines, respectively. Numbers 1–4 indicate the in situ measured peak concentrations above  $200 \mu\text{g}/\text{m}^3$  shown in Figure 3a. (top) Retrieved ash columns from SEVIRI, followed by the modeled ash clouds as simulated by FLEXPART run on ECMWF meteorology with a priori emissions from Figure 1, and emissions from a simple source term assuming a uniform vertical distribution. (bottom) Modeled total ash columns from the three simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM) initialized with the a posteriori source terms of Figure 1.

comparison. Thereafter, a comparison of observed and modeled PM<sub>10</sub> concentrations and particle size distributions at the Jungfraujoch station is given. Finally, the detailed statistical model-observation analysis of the whole in situ measurement data set from all BAe-146 and Falcon flights performed in April and May, the DIMO flight and the Jungfraujoch observations, is presented.

### 3.2.1. Aircraft Measurements

#### 3.2.1.1. Case Study 14 May

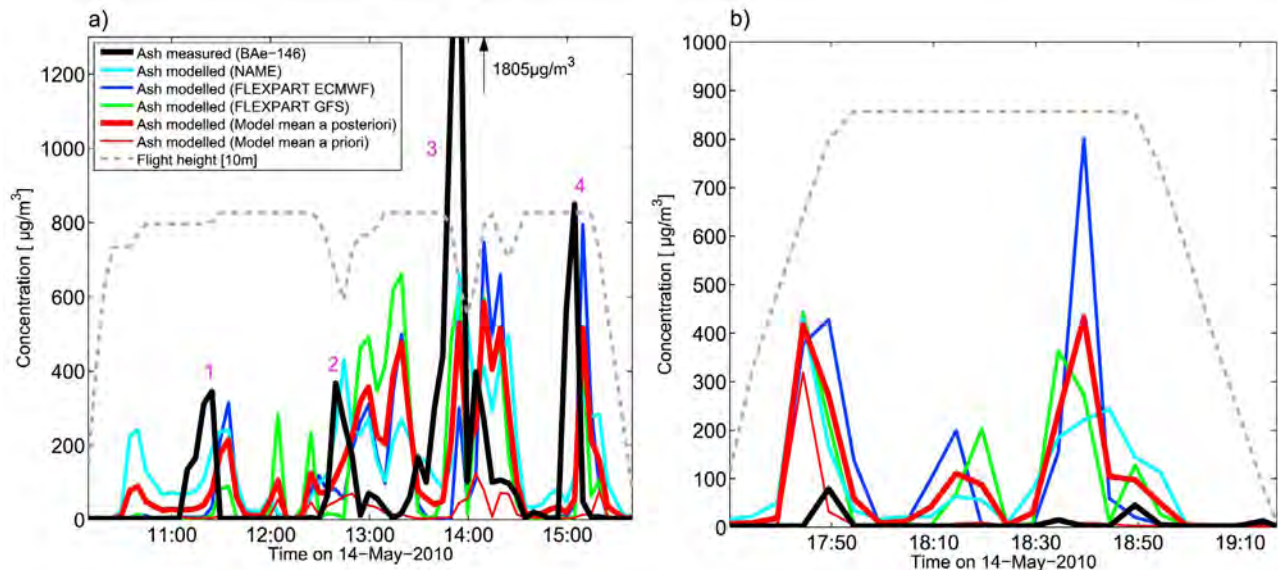
[46] On 14 May 2010 a highly concentrated ash cloud was located over Northern England. The SEVIRI satellite instrument observed an ash cloud between 11:00 and 12:00 UTC (Figure 2) with maximum column loadings around  $54\text{--}56^\circ\text{N}$ . The a priori and uniform model simulations do not have the ash cloud far enough south, while the posteriori model simulations (lower panel) capture the observed ash cloud better. This is expected as the SEVIRI data was used to estimate the a posteriori source term. Also *Devenish et al.* [2012] show that the ash cloud over UK on 14 May is particularly sensitive to the source profile. However, our a posteriori simulations show lower total column values than SEVIRI, but improvements in the satellite retrieval scheme

suggest the satellite values are decreased by 20–30% to what is presented here (Prata and Prata, submitted manuscript, 2011, Figure 13).

[47] The BAe-146 research aircraft performed two flights over UK on 14 May and took measurements from 10:06 UTC to 15:41 UTC and in the evening between 17:28 UTC and 19:17 UTC. Figure 2 shows the flight tracks while Figure 3 shows the measured in situ ash concentrations for the two flights on this day (black lines). A screening of the measurements for volcanic ash has been made (Johnson et al., submitted manuscript, 2011), and there are four main detections of the ash cloud during the first flight (peaks denoted by numbers 1–4).

[48] The three different a posteriori model simulations (green, blue and turquoise lines in Figure 3a) all have similar ash concentration peaks as observed. The modeled ash concentrations are mostly within a factor of two of the in situ measurements, and thus within the uncertainty of the observations. FLEXPART generally simulates higher concentrations than NAME, but both models show the peaks in the same time-interval. However, it seems there is a shift in the





**Figure 3.** In situ measured ash concentrations (black lines) and modeled ash concentrations from different simulations (turquoise lines: NAME-MetUM a posteriori, blue lines: FLEXPART-ECMWF a posteriori, green lines: FLEXPART-GFS a posteriori, red thick lines: Model mean a posteriori, red thin lines: Model mean a priori) along the BAe-146 flight tracks over the UK on 14 May 2010 between (a) 10:06–15:41 UTC (Flight 528) and (b) 17:28–19:17 UTC (flight 528b). Numbers 1–4 indicate the peak concentrations whose locations are marked in Figure 2. The measured peak concentration around 13:50 UTC may include a contribution from ice or ice coatings leading to a possible overestimate of mass concentration for that profile (Johnson et al., submitted manuscript, 2011). Flight height is shown by the gray dashed line (in units of 10 m).

modeled peaks to the observations most likely due to position errors of the modeled ash clouds.

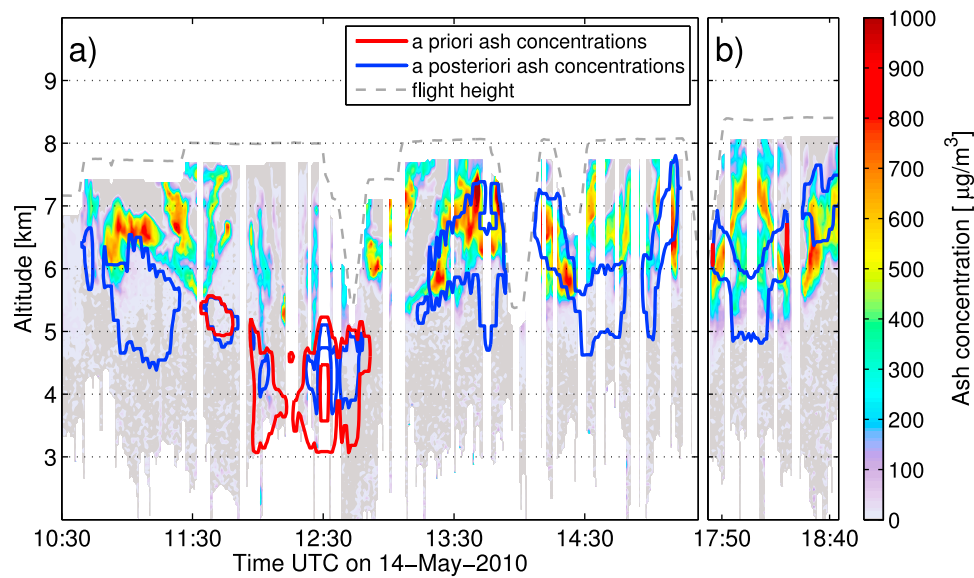
[49] The mean ash concentrations over the three a posteriori simulations (red thick line) are generally much higher than the mean ash concentrations over the three a priori simulations (red thin line) and are in more agreement with the measured concentrations. The simulation assuming a uniform source profile does not give any ash signal over these flight tracks (not shown). The observed ash cloud was emitted from the volcano 1–2 days prior to the observation time and the larger a posteriori concentrations are related to the clear increase of the a posteriori ash emissions over the a priori emissions between 12 and 13 May (Figure 1). For the second flight (Figure 3b) the modeled ash concentrations are overestimated compared to the measurements, and are related to large ash emissions early on 14 May.

[50] The highest concentrations observed during the whole event from the BAe-146 aircraft were observed over Scotland around 13:50 UTC. There is some evidence of ice or ice coatings leading to a possible overestimate of the ash mass during this profile [Marenco et al., 2011; Johnson et al., submitted manuscript, 2011]. The measured maximum concentration was  $\sim 1800 \mu\text{g}/\text{m}^3$  (five minute average). The modeled maximum concentration for this peak is  $580 \mu\text{g}/\text{m}^3$  for the mean over the a posteriori simulations. Thus, the models did not capture single in situ observed peak values on the order of  $2000 \mu\text{g}/\text{m}^3$ .

[51] The lidar observations from the BAe-146 aircraft on 14 May show ash layers at 6–7 km with little ash below 5 km (Figure 4). The vertical position of the a posteriori modeled

ash cloud (blue line) seems to be slightly too low compared to the observations, but agrees within  $\pm 1$  km. The concentrations are slightly overestimated but within the measurement uncertainty of a factor of two. There is a much better agreement with the observations for the a posteriori over the a priori ash cloud (red line). The maximum airborne lidar retrieved ash concentration for the flights on 14 May was  $1900 \mu\text{g}/\text{m}^3$  around 13:40 UTC at 7.3 km [Marenco et al., 2011], in accordance with the maximum in situ measurement. The modeled a posteriori concentration at about the same time and altitude is  $\sim 2100 \mu\text{g}/\text{m}^3$ .

[52] The modeled ash cloud around 12:30 UTC at about 4 km, which seems rather robust between the simulations, is difficult to evaluate because the lidar data are heavily screened in this part due to clouds (white areas indicate no data). However, there are indications that the modeled ash cloud concentrations are overestimated in this area and that the real ash layer was located at a higher altitude. This discrepancy coincides with peak 2 of the in situ measurements where the models did not perform that well (Figure 3). Also SEVIRI detected no ash at the location of peak 2 around 12:30 UTC. The satellite instrument's detection limit is  $\sim 200 \mu\text{g}/\text{m}^3$  (Prata and Prata, submitted manuscript, 2011) and thus the in situ measured concentration of  $\sim 350$ – $400 \mu\text{g}/\text{m}^3$  at peak 2 should be detectable by the satellite. Other satellite data (e.g., MODIS) suggests there were clouds at this location, but it is thought not to be sufficient to cause problems for the satellite retrieval. SEVIRI detected ash at the location of peak 2 about an hour earlier (Figure 2), and the ash cloud was moving quickly eastward, thus an ash



**Figure 4.** Comparisons of measured and modeled lidar profiles along the BAe-146 flight tracks on 14 May 2010 between (a) 10:30–15:30 UTC (Flight b528) and (b) 17:30–19:00 UTC (flight b528b). The estimated ash mass concentrations from the lidar signal are overlaid by isolines of a priori and a posteriori modeled ash mass concentrations at  $500 \mu\text{g}/\text{m}^3$  which represent the mean concentrations over the three model simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM). Flight height (km) is shown by the gray dashed line.

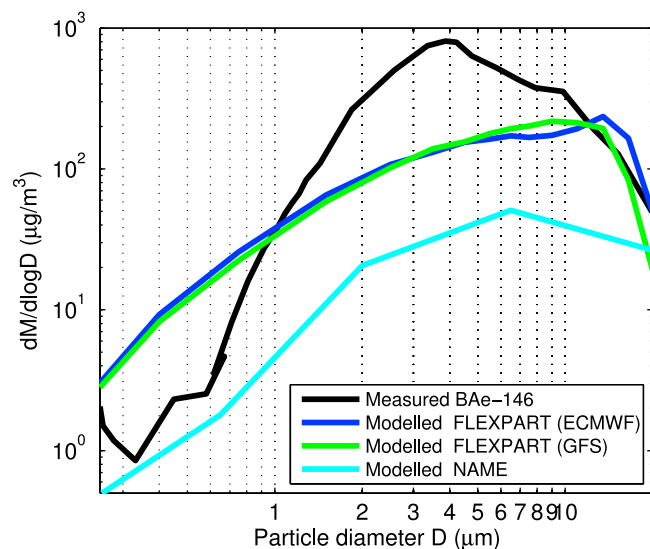
cloud to the west at a later time seems unlikely. This illustrates the uncertainties related to the measurement-model comparison at this particular time.

[53] The particle size distribution as measured from the BAe-146 flights on 14 May is shown in Figure 5 together with the modeled size distribution from the FLEXPART ECMWF-based and GFS-based a posteriori model simulations, and a NAME particle size distribution. The particle size distributions represent the average distribution over all the BAe-146 locations where there was an observed and confirmed ash signal during the flights.

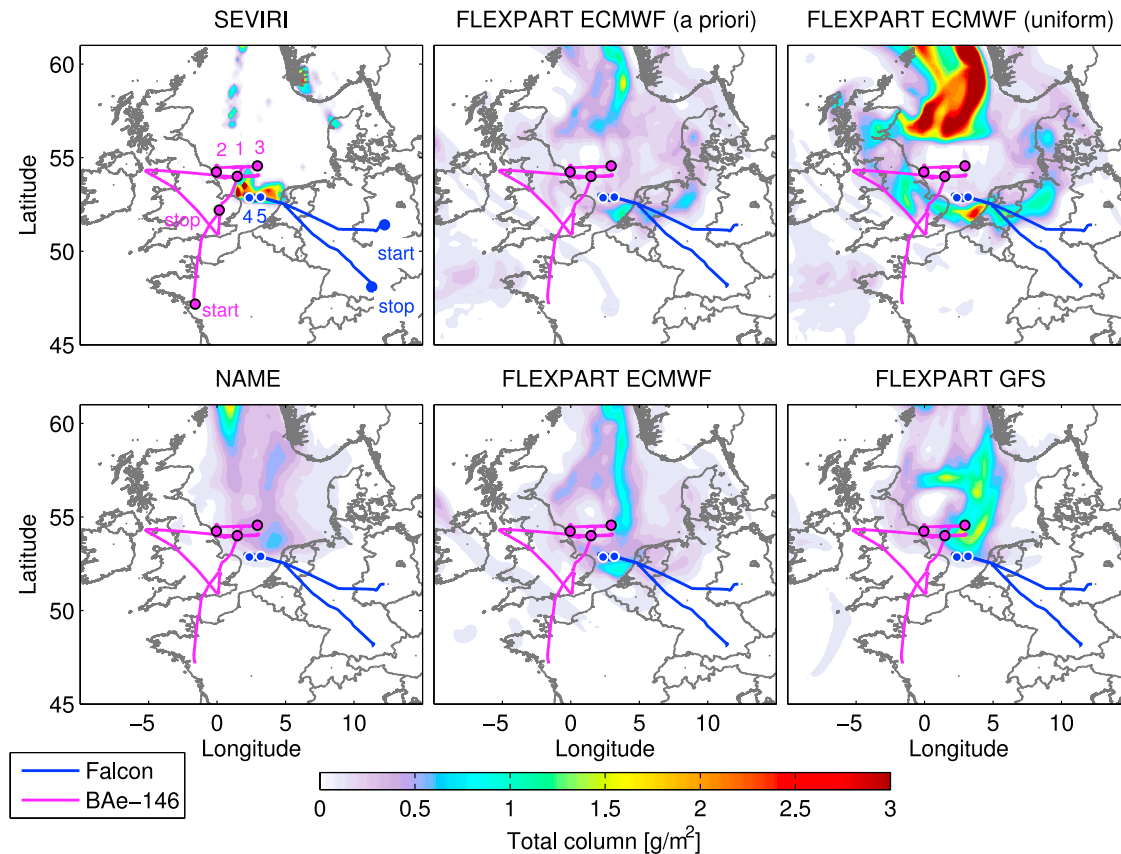
[54] The modeled particle size distribution clearly depends on the chosen initial particle size distribution at the source and on the particle density. The NAME size distribution is retrieved from simulations using a different source term, initial particle size distribution and set-up [Dacre *et al.*, 2011] as the simulations presented in this paper. Thus, the NAME and FLEXPART particle size distributions cannot be directly compared, but it allows an evaluation of the shape of the various modeled particle size distributions after long-range transport when sedimentation and other removal processes have taken place. Only the coarse-mode (particle diameter  $>0.25 \mu\text{m}$ ) of the measured size distribution is simulated by the model, and thus only this size range is shown.

[55] Comparing the relative shapes of the in situ measured and modeled particle size distributions indicates that the model has too little mass in the  $1\text{--}10 \mu\text{m}$  size range relative to larger particle sizes. The measured size distribution is dominated by particles around  $4 \mu\text{m}$  diameter, while the FLEXPART modeled size distributions have a peak around  $10 \mu\text{m}$ , and the NAME distribution peaks at  $6.5 \mu\text{m}$ . Thus, the modeled distributions seem to be shifted to larger particle

sizes, i.e., the peak in the modeled size distributions is found at a larger particle size than the peak in the in situ measured size distribution. However, Prata and Prata (submitted manuscript, 2011) found mean effective particle diameters of  $8\text{--}12 \mu\text{m}$  from the SEVIRI satellite data. Overall, the in situ measured and modeled particle size distributions compare reasonably well for this case study.



**Figure 5.** Measured and modeled particle size distributions for the peak ash concentrations from Figure 3 measured during the BAe-146 flight on 14 May 2010. An ash particle density of  $2.3 \text{ g}/\text{cm}^3$  is assumed for the volume to mass conversions.



**Figure 6.** Observed and modeled total ash columns on 17 May 2010 between 16:00–17:00 UTC. The flight tracks for the Falcon flight and the BAe-146 flight (b530) are shown as blue and magenta lines, respectively. Numbers 1–5 indicate the in situ measured peak concentrations above  $100 \mu\text{g}/\text{m}^3$  shown in Figure 7. (top) Retrieved ash columns from SEVIRI, followed by the modeled ash cloud as simulated by FLEXPART run on ECMWF meteorology with the a priori emissions from Figure 1, and emissions from a simple source-term assuming a uniform vertical distribution. (bottom) The modeled total ash columns from the three simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM) initialized with the a posteriori source terms of Figure 1.

### 3.2.1.2. Case Study 17 May

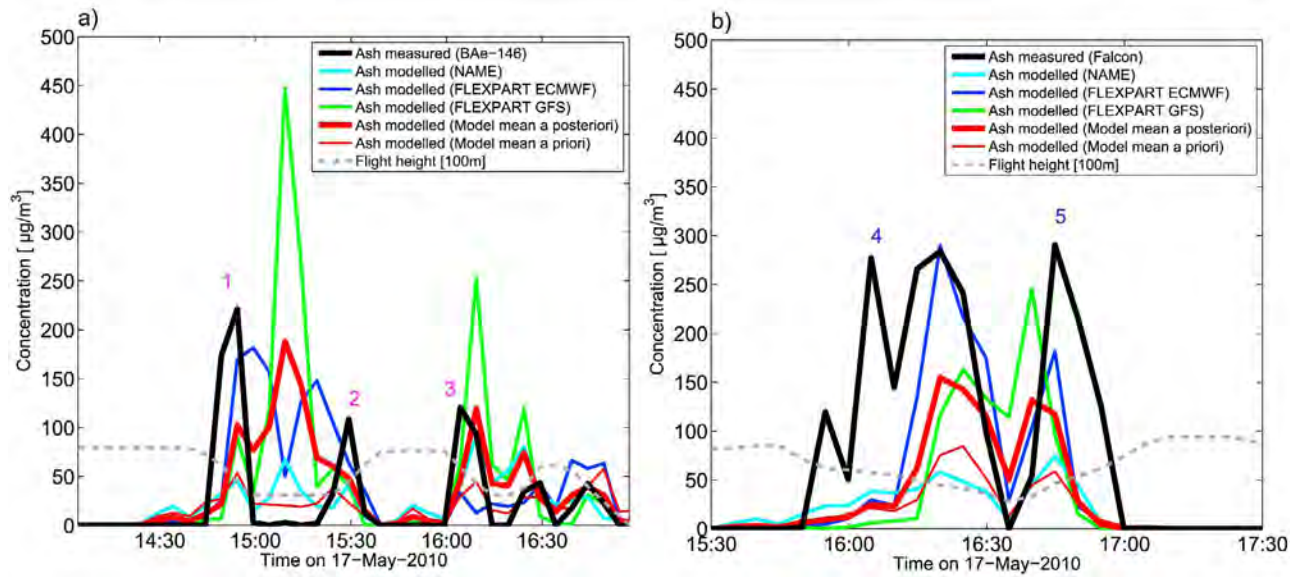
[56] On 17 May 2010 the BAe-146 and the Falcon aircraft flew through an ash cloud over the North Sea. The SEVIRI satellite instrument observed an ash cloud between the Netherlands and UK, while the modeled ash clouds are broader and cover a larger part of the North Sea (Figure 6). The three a posteriori simulations (lower panel) show rather different structures of the ash cloud covering the North Sea, but they reduce the anomalous ash filament over the UK that was predicted in real-time by the London VAAC [Turnbull *et al.*, 2012, Figure 2] and that is also seen in particular in the uniform simulation, but also in the a priori simulation. The measurements taken by BAe-146 over South-East England and other parts of the UK showed that these areas were free of detectable ash which confirmed a decision to re-open the low-level airspace on that day (Johnson *et al.*, submitted manuscript, 2011). In this case study, the FLEXPART ECMWF-based a posteriori simulation seems to be in better agreement with the satellite observations. This simulation especially captures the thin filament of the ash cloud along the coast of the Netherlands which FLEXPART-GFS and NAME do not simulate as well probably due to the meteorology driving the models. The flight tracks in Figure 6

show that this was the ash cloud that the Falcon aircraft was targeting.

[57] For the BAe-146 observations from 14:30 to 15:45 UTC (Figure 7a, peaks denoted 1 and 2) the differences between the models (green, blue, turquoise lines) are rather large and the timing of the measured and modeled peaks is not closely related. The mean a posteriori concentrations (thick red line) is within the factor two uncertainty of the observations (black line), but the modeled peak falls between the two measured peaks (peak 1 and 2). The a priori concentrations are again below the uncertainty range of the observations. The FLEXPART ECMWF-based a posteriori simulation is also in better agreement with these in situ measurements, showing a double peak in this time period, and concentrations in closest agreement with the observations.

[58] Between 16:00–17:00 UTC the two aircraft took simultaneous in situ measurements of the ash cloud; BAe-146 measured five-minute average ash concentrations up to  $120 \mu\text{g}/\text{m}^3$  (Figure 7a, peak 3) while Falcon measured higher concentrations up to  $300 \mu\text{g}/\text{m}^3$  (Figure 7b, peak 4 and 5). A comparison of the observations from both aircraft is presented in detail by Turnbull *et al.* [2012]. The





**Figure 7.** In situ measured ash concentrations (black lines) and modeled ash concentrations from different simulations (turquoise lines: NAME-MetUM a posteriori, blue lines: FLEXPART-ECMWF a posteriori, green lines: FLEXPART-GFS a posteriori, red thick lines: Model mean a posteriori, red thin lines: Model mean a priori) along the flight tracks on 17 May 2010 for (a) the BAe-146 flight (Flight B530) and (b) the Falcon flight. Numbers 1–5 indicate the peak concentrations whose locations are marked in Figure 6. Flight height is shown by the gray dashed line (in units of 100 m).

differences in the observations are mainly due to the fact that the aircraft clearly did not sample the exact same regions of the ash clouds, but are also due to slightly different instrumentation and assumptions for the post processing of the measurement data. The main peaks of the observations are captured by the models, and a posteriori concentrations are generally within the uncertainty of the measurements, while the a priori values are again lower. The better agreement for FLEXPART-ECMWF to the in situ observations confirms the presence of the ash filament along the Netherlands as observed also by SEVIRI (Figure 6).

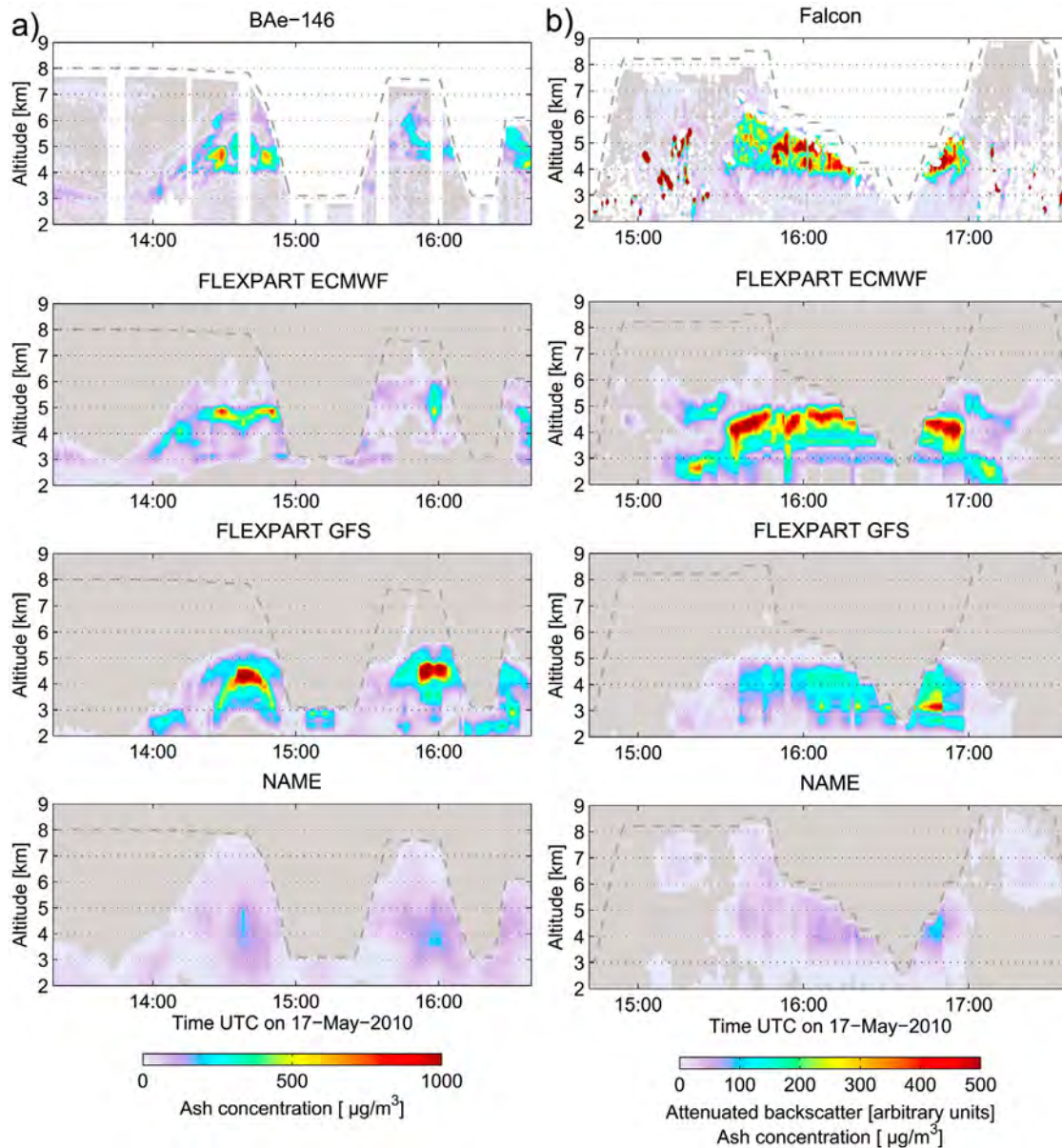
[59] Both the BAe-146 and the Falcon lidars observed an ash cloud centered around 4.5 km altitude on 17 May (Figure 8). Note that the patchy signal in the Falcon lidar profiles from departure up to about 15:20 and from 17:00 is due to clouds (and not ash) located over the Netherlands and Germany [Schumann *et al.*, 2011, Figure 16]. There are rather large differences between the modeled lidar profiles from the three a posteriori model simulations. All simulations show an ash signal, but the maximum values and positioning of the ash cloud is quite different. The NAME simulation has a much broader extent of the ash cloud with lower peak concentrations than FLEXPART. The difference between FLEXPART and NAME is probably due to the representations of subgrid-scale diffusion (unresolved eddies) in the models. Devenish *et al.* [2012] have investigated the sub-grid diffusion with NAME for the 14 May ash cloud over UK and found that the vertical extent of the ash cloud was much reduced when no vertical diffusion or horizontal meander were applied in the model. Other model parameterizations like deposition and sedimentation may also contribute to the differences between the models. Visually the FLEXPART simulations, in particular the ECMWF-based simulation, look more similar to the lidar observations because they

reproduce the patchy and inhomogeneous character better. However, the patches are not necessarily always in the right place. The NAME model doesn't seem to get the position of ash cloud any better, but because the clouds are broader and more widely spread they seem to encompass the locations of observed patches more. Therefore, the conclusions of which model is "better" may depend on the goal of the simulations and the relative costs of misses (failure to predict an ash layer that is there in reality) versus false alarms.

[60] The measured and modeled particle size distributions for 17 May are shown in Figure 9. As for the 14 May case study, the BAe-146 size distribution peaks around  $4\ \mu\text{m}$  (Figure 9a), while the modeled size distributions show too little mass in the  $1\text{--}10\ \mu\text{m}$  size range when comparing the relative shapes of the modeled and the in situ measured size distributions. Note that the peak in the measured size distribution at small particle sizes (diameter  $<0.5\ \mu\text{m}$ ) is ascribed to secondary aerosols (e.g., sulfate) which the models do not simulate.

[61] The size distribution measured by the Falcon (Figure 9b) is significantly different from that measured by BAe-146 and is dominated by particles in the range of  $10\text{--}20\ \mu\text{m}$ . The shape of the Falcon size distribution is discussed in detail by Schumann *et al.* [2011], and the differences between the Falcon and BAe-146 measured distributions are examined in detail by Turnbull *et al.* [2012]. The main source of difference is the assumptions of the particle properties (particle shape and the degree of absorption applied to the particle refractive index), but also uncertainties in e.g., the calibration between the instruments.

[62] The shape of the modeled size distributions are in closer agreement with the Falcon distribution than with the BAe-146 distribution. This is maybe not surprising as a Falcon size distribution (from the 2 May flight) was used as



**Figure 8.** Comparisons of measured and modeled lidar profiles from the three a posteriori model simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM) along the BAe-146 and Falcon flight tracks on 17 May 2010. (a) Estimated ash mass concentrations from the BAe-146 lidar signal and (b) attenuated backscatter observed during the Falcon flight. White areas within the lidar data indicate missing data. The two ash concentration color scales apply to the modeled lidar profiles on the left and right panel, respectively. Flight height (km) is shown by the gray dashed line.

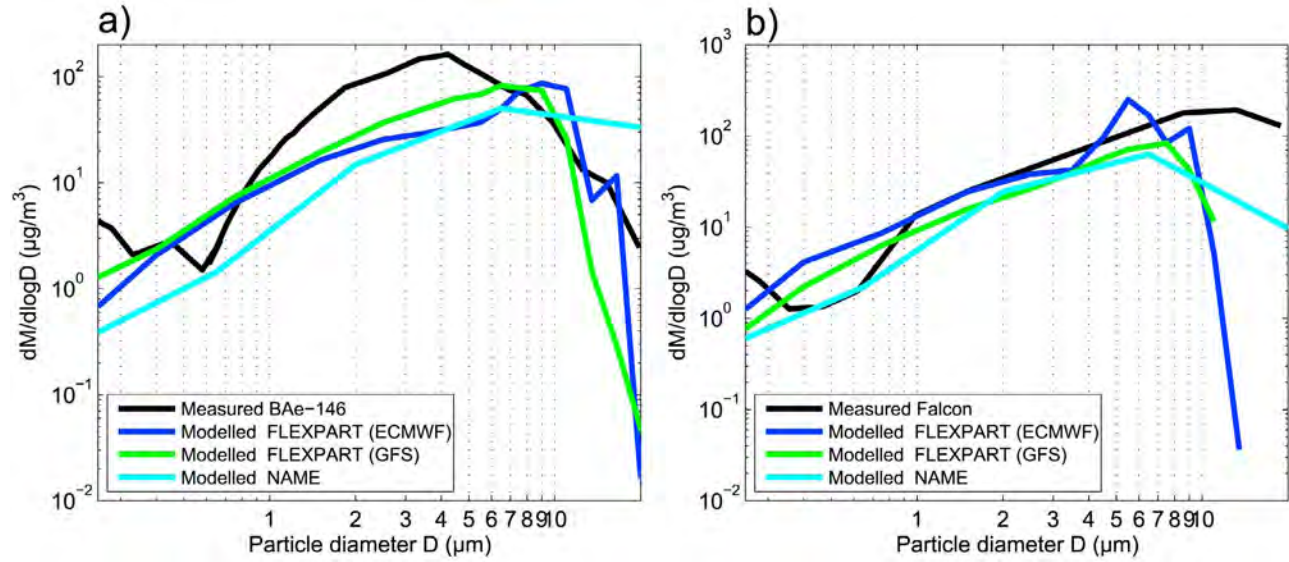
guidance for determining the initial size distribution used for the FLEXPART model simulations. For the Falcon size distribution there is a too quick drop-off for large particles in the FLEXPART model compared to the measured distribution.

### 3.2.1.3. Case Study 18 May

[63] On 18 May 2010 the ash cloud was transported further southwards and stretched from the North Sea down to the Alps. On this day, aircraft measurements were performed by three different aircraft; BAe-146, Falcon and the Swiss DIMO. Figure 10 shows the a posteriori modeled ash clouds and the flight tracks of the three flights. The BAe-146 observed the ash cloud in about the same location as on

17 May with similar results but lower ash concentrations. The Falcon flew over Germany and the North Sea and observed multiple ash layers during almost the entire flight. The DIMO aircraft sampled the ash cloud over the Swiss Alps in an area with lower total columns of ash. The SEVIRI satellite instrument detected mass loading in the regions of the aircraft measurements around or below  $1 \text{ g/m}^2$  (not shown). All ash concentrations below  $200 \text{ µg/m}^3$  are unlikely to have been detected by SEVIRI. It should be noted that the mass loading from SEVIRI are larger and more widespread about two hours earlier than the measurements are taken.



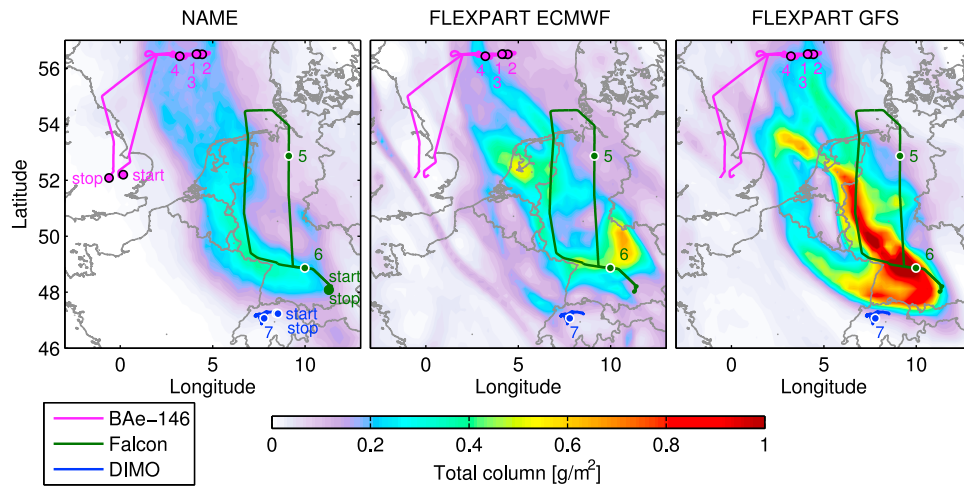


**Figure 9.** Measured and modeled particle size distributions for the peak ash concentrations from Figure 7 measured during (a) the BAE-146 flight and (b) the Falcon flight on 17 May 2010. An ash particle density of  $2.3 \text{ g/cm}^3$  is assumed for the volume to mass conversions for both BAE-146 and Falcon.

[64] The models perform overall very well when comparing the a posteriori concentrations to the measurements (Figure 11). All the measured ash concentrations are below  $200 \text{ } \mu\text{g/m}^3$  and thus did not pose any hazard to aviation, but are of academic interest for the model-measurement comparisons. The four ash concentrations peaks measured by BAE-146 are all simulated by the models (Figure 11a), but as opposed to the other case studies presented, the models slightly overestimate the observed concentrations, and the a posteriori values are reduced compared to the a priori values. The measured maximum concentration from BAE-146 reached  $50 \text{ } \mu\text{g/m}^3$  (peak 3) (five minute average), and the modeled concentrations are within the factor two uncertainty

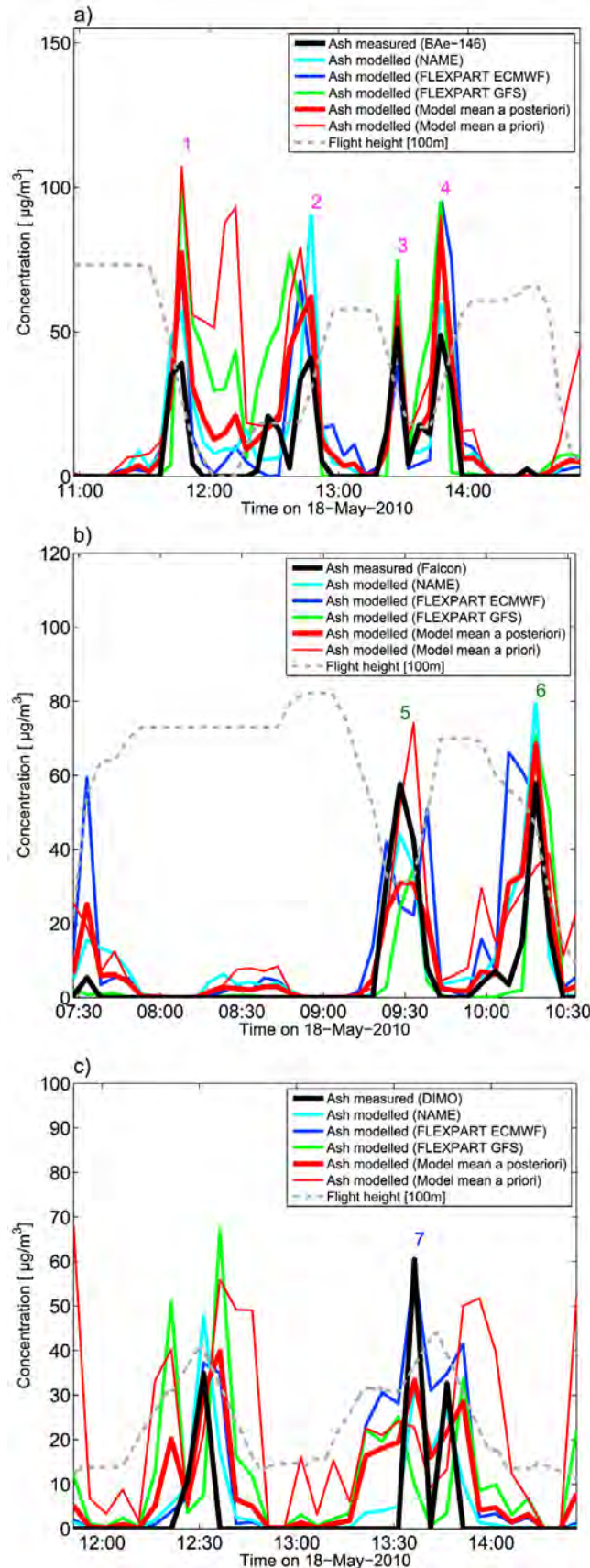
of the measurements. The Falcon measured a maximum concentration of  $60 \text{ } \mu\text{g/m}^3$  (five minute average) over Southern Germany (Figure 11b, peak 6) and the modeled concentrations are in good agreement with the measured ones. The observed peaks by DIMO are also captured by the models and especially well with the NAME and FLEXPART-ECMWF a posteriori simulations.

[65] In general the fit between the modeled ash concentrations and the DIMO observations is not as good as for the two other flights. The DIMO flew further south, on the edge of the ash cloud which is likely to lead to concentrations being less predictable and sensitive to small differences in the modeling. Also, the complex topography of this region



**Figure 10.** Modeled total ash columns on 18 May 2010 between 12:00–13:00 UTC from the three simulations (NAME-MetUM, FLEXPART-ECMWF and FLEXPART-GFS) initialized with the a posteriori source term of Figure 1. The flight tracks for the BAE-146 flight (b531), the Falcon flight and the DIMO flight are shown as magenta, green and blue lines, respectively. Numbers 1–7 indicate the in situ measured peak concentrations shown in Figure 11.





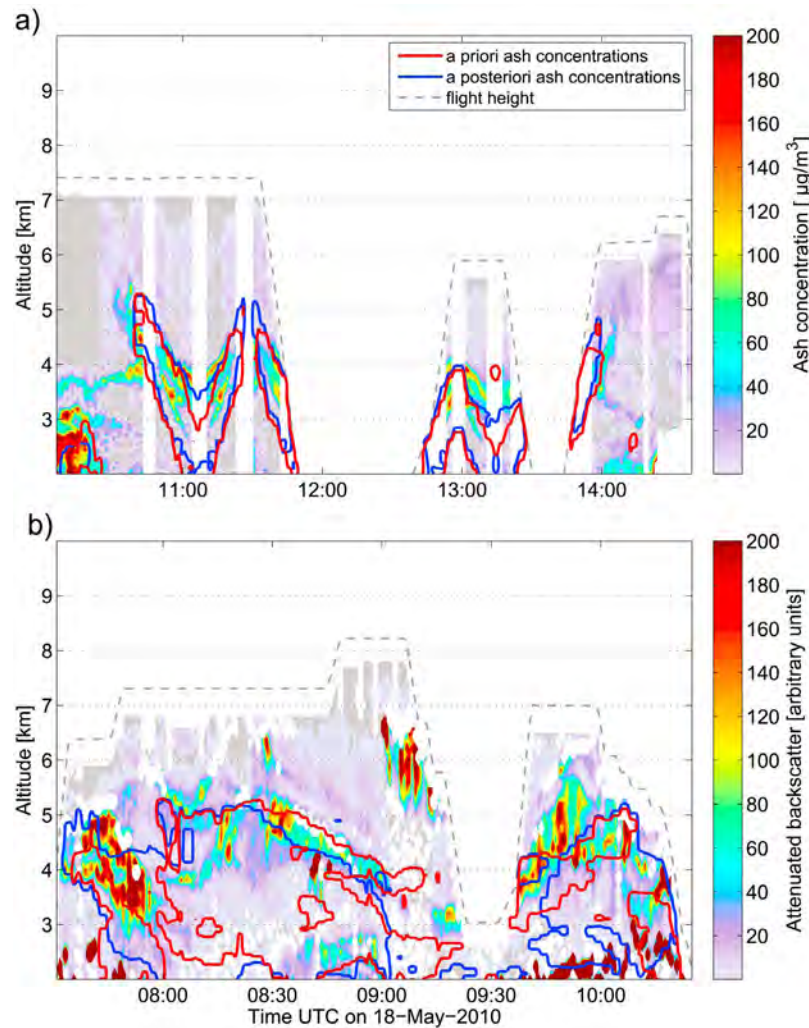
makes it harder for the models to get the timing and position of the ash cloud right. Furthermore, the DIMO flight covered an area much smaller than the BAE-146 and Falcon, which gives less contrast in the modeled ash clouds with the specified model resolution. Also, as noted before, SEVIRI detected larger mass loading and a more widespread ash cloud two hours previous to the measurements, which might suggest there is a slight timing error in the modeled ash cloud.

[66] The lidar profiles from the aircraft on 18 May reveal ash clouds at various heights (Figure 12). The shape and mean concentrations of the modeled ash clouds agree very well with the BAE-146 lidar observations (Figure 12a). There are no significant differences for the mean over the three a priori and the three a posteriori simulations for this case. The uniform simulation (not shown) has much higher concentrations than observed, but the structure is well modeled.

[67] The lidar profiles from the Falcon are difficult to interpret because the atmospheric layering was quite complex. However, the modeled ash cloud is still in fair agreement to the lidar signal in that the timing of the modeled ash cloud corresponds roughly to when the lidar showed strong signals. Similar to the BAE-146 lidar profile, the a priori and a posteriori ash clouds are not significantly different. It seems that the model simulations have too much ash at low altitudes (below 3–4 km) where the lidar shows no signal. It should be noted that there are large differences between the three different a posteriori model simulations (not shown) for this lidar comparison, with the FLEXPART-GFS simulation showing the strongest ash signal at lower altitudes. Note also that the strong lidar signal at 5–6 km around 09:00 to 09:15 is caused by clouds and not an ash cloud as the signal below is strongly attenuated indicating high extinction by clouds above. In addition, satellite imagery showed the presence of clouds over the North Sea where the measurements were taken [see Schumann *et al.*, 2011, Figure 18], and the in situ measurements showed no presence of ash at these altitudes (Figure 11b).

[68] The comparison of the shape of the measured and modeled particle size distributions shown in Figure 13 suggests a fairly good agreement, but as for the 17 May case the models have too little mass in the 1–10  $\mu\text{m}$  size range compared to BAE-146, and FLEXPART has a too quick drop-off for larger particles compared to the Falcon distribution. The particle size distribution for the DIMO flight peaks at 2.5  $\mu\text{m}$ , thus the peak is shifted to smaller particles as the plume is transported further south and the large particles are lost due to sedimentation. The FLEXPART-GFS particle size distribution is in overall good agreement with the DIMO distribution, while the FLEXPART-ECMWF is shifted to larger particles.

**Figure 11.** In situ measured ash concentrations (black lines) and modeled ash concentrations from different simulations (turquoise lines: NAME-MetUM a posteriori, blue lines: FLEXPART-ECMWF a posteriori, green lines: FLEXPART-GFS a posteriori, red thick lines: Model mean a posteriori, red thin lines: Model mean a priori) along the flight tracks on 18 May 2010 for (a) the BAE-146 flight (Flight b531), (b) the Falcon flight and (c) the DIMO flight. Numbers 1–7 indicate the peak concentrations whose locations are marked in Figure 10. Flight height is shown by the gray dashed line (in units of 100 m).



**Figure 12.** Comparisons of measured and modeled lidar profiles along the BAe-146 and Falcon flight tracks on 18 May 2010. (a) Estimated ash mass concentrations from the BAe-146 lidar signal and (b) attenuated backscatter observed during the Falcon flight. The measured values are overlaid by isolines of a priori and a posteriori modeled ash mass concentrations of  $50 \mu\text{g}/\text{m}^3$  which represent the mean values over the three model simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM). White areas within the lidar data indicate missing data. Flight height (km) is shown by the gray dashed line.

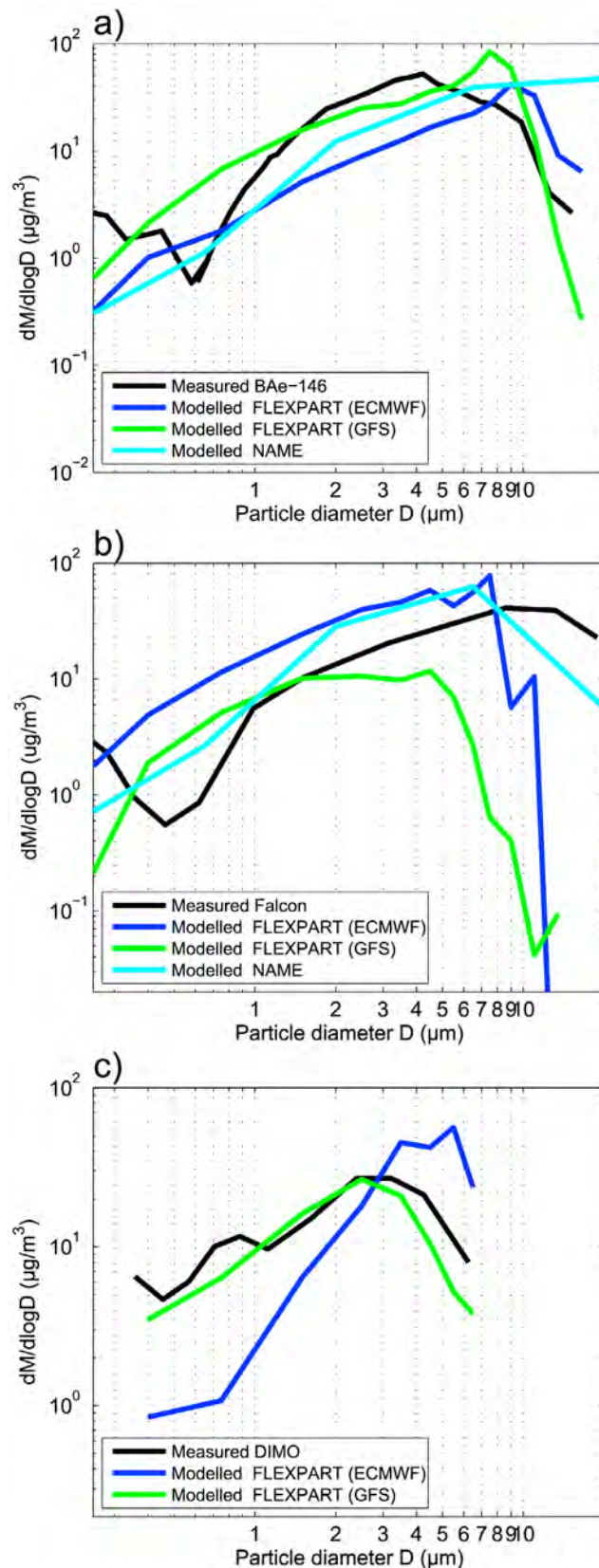
### 3.2.2. Surface Measurements at Jungfraujoch

[69] Two episodes with increased PM<sub>10</sub> concentrations were related to volcanic aerosol clouds over Jungfraujoch in April and May 2010 [Bukowiecki *et al.*, 2011]. Modeled PM<sub>10</sub> concentrations from the FLEXPART simulations are estimated by summing the mass concentrations from modeled ash particle sizes  $10 \mu\text{m}$  diameter and smaller. The NAME model results include the whole particle size distribution spectrum for the model simulations ( $0.25\text{--}250 \mu\text{m}$ ) because of the model output definitions, i.e., for NAME the ash concentrations are not PM<sub>10</sub>. However, as the FLEXPART modeled size distributions suggest particles up to  $10\text{--}11 \mu\text{m}$ , and assuming that about the same particle size distribution applies to NAME, the results may be considered to roughly correspond to PM<sub>10</sub> concentrations. This suggests that most of the particles larger than  $10\text{--}11 \mu\text{m}$  had fallen out when the ash cloud reached the Swiss Alps.

[70] Both the measured and modeled PM<sub>10</sub> concentrations are hourly average values and the modeled concentrations are extracted from the vertical model layer where the true altitude of the station (3580 m a.s.l.) is found. The topography at the Jungfraujoch location in the models is only  $\sim 1600\text{--}1800$  m a.s.l. (FLEXPART ECMWF: 1780 m a.s.l., FLEXPART GFS: 1615 m a.s.l., NAME: 1656 m a.s.l.), because of the low model resolution. Thus, there are actually eight model layers between the model ground and the true measurement altitude. Extracting the model concentrations from the model layer of the true altitude of the station is thought as most reasonable, however it introduces some uncertainty.

[71] In general, there are difficulties of simulating transport to a high altitude station, as seen in previous studies [De Wekker *et al.*, 2004; Weigel *et al.*, 2007; Stohl *et al.*, 2009]. To fully capture the complex flow structures that develop in the complex terrain of the Alps, meteorological



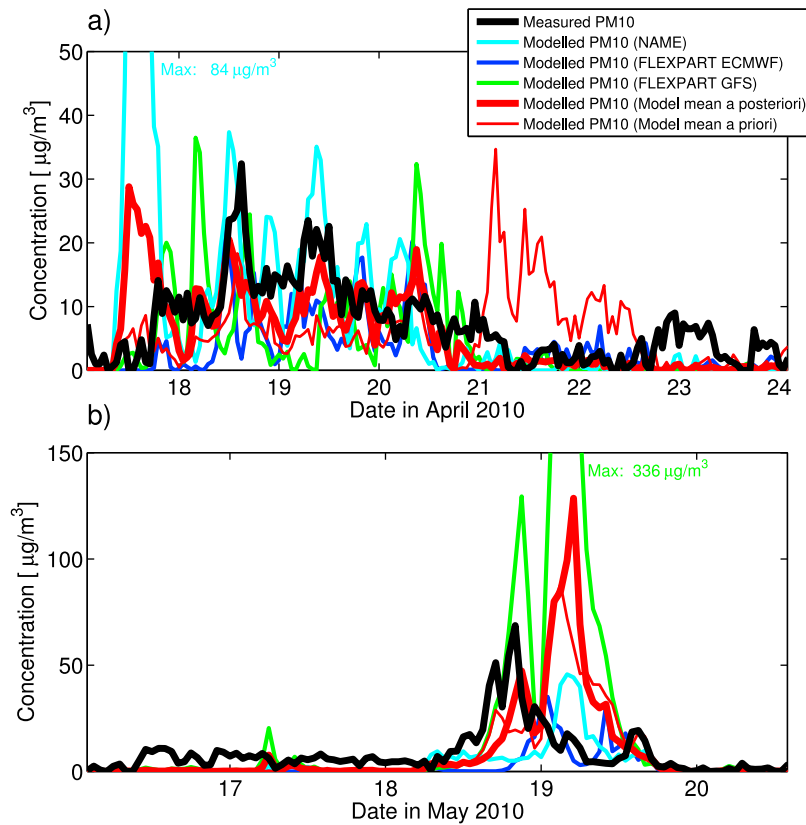


models with grid sizes around and below 1 km need to be applied [e.g., *Monks et al.*, 2009]. This is beyond the current capabilities of the ECMWF, GFS and the MetUM weather forecast models. Therefore, we expect that model uncertainties in complex terrain will hamper the comparison with observations as compared to free tropospheric and flat terrain observations. These uncertainties are further amplified for a volcanic ash cloud that has been transported over a long distance before reaching the measurement station. During the flights with the DIMO aircraft over the Swiss Alps it was seen that the ash was transported into the valleys and was thereafter during the diurnal cycle lifted to the higher altitudes of the Jungfraujoch station. Thus, the ash was not transported directly to the observation station at high altitudes. This complex transport pattern is difficult for the models to capture.

[72] The observed PM<sub>10</sub> concentrations are compared with modeled PM<sub>10</sub> concentrations in Figure 14. In general, the models produce similar ash concentrations peaks as observed at the station. There is an improvement in the agreement to the measurements for the a posteriori simulations over the a priori simulations. In particular, there is a significant reduction in the modeled PM<sub>10</sub> concentrations for the a posteriori-mean over the a priori-mean on 21–22 April (Figure 14a). This is related to the large differences in the a priori and a posteriori ash emissions around 17 April (Figure 1). The better fit to the observations for the a posteriori in this time period and for this particular location suggests that the reduced a posteriori emissions around 17 April seen in Figure 1 were realistic.

[73] For the May event (Figure 14b), the models also produce an ash signal over Jungfraujoch. The magnitude of the observed PM<sub>10</sub> concentration peak is simulated quite well, but it seems like the ash cloud is modeled to arrive at the station about 12 h later than the observed PM<sub>10</sub> concentration peak. The time delay was found in all the eight model layers from the model ground up to the true altitude of the station. This indicates that the modeled transport is too slow toward the Alps, and thus the real ash cloud was extending further south, which was also indicated in the study by *Bukowiecki et al.* [2011] and also in section 3.2.1.3 where the SEVIRI observations indicate that the models were moving the ash cloud too slowly. Furthermore, there are large differences for the three a posteriori simulations which are due to a very narrow plume transported across the Alps simulated differently with the models and the three different meteorological data sets (not shown). Thus, the distribution of the ash was quite inhomogeneous making it difficult for the models to correctly capture the fine structures. This was also seen for the model comparisons to the DIMO measurements in section 3.2.1.3. Considering the difficulties in simulating transport to a high altitude station, the results are encouraging.

**Figure 13.** Measured and modeled particle size distributions for the peak ash concentrations from Figure 11 measured during (a) the BAe-146 flight (b) the Falcon flight, and (c) the DIMO flight on 18 May 2010. An ash particle density of  $2.3 \text{ g}/\text{cm}^3$  is assumed for the volume to mass conversions for both BAe-146 and Falcon, while  $2.65 \text{ g}/\text{cm}^3$  is used for DIMO.



**Figure 14.** Measured PM10 concentrations (black lines) and modeled PM10 concentrations from different simulations (turquoise lines: NAME-MetUM, blue lines: FLEXPART-ECMWF, green lines: FLEXPART-GFS, red thick lines: Model mean a posteriori, red thin line: Model mean a priori) at the Jungfraujoch station (46.55°N, 7.99°E, 3580 m a.s.l.) for (a) 17–23 April 2010 and (b) 16–20 May 2010.

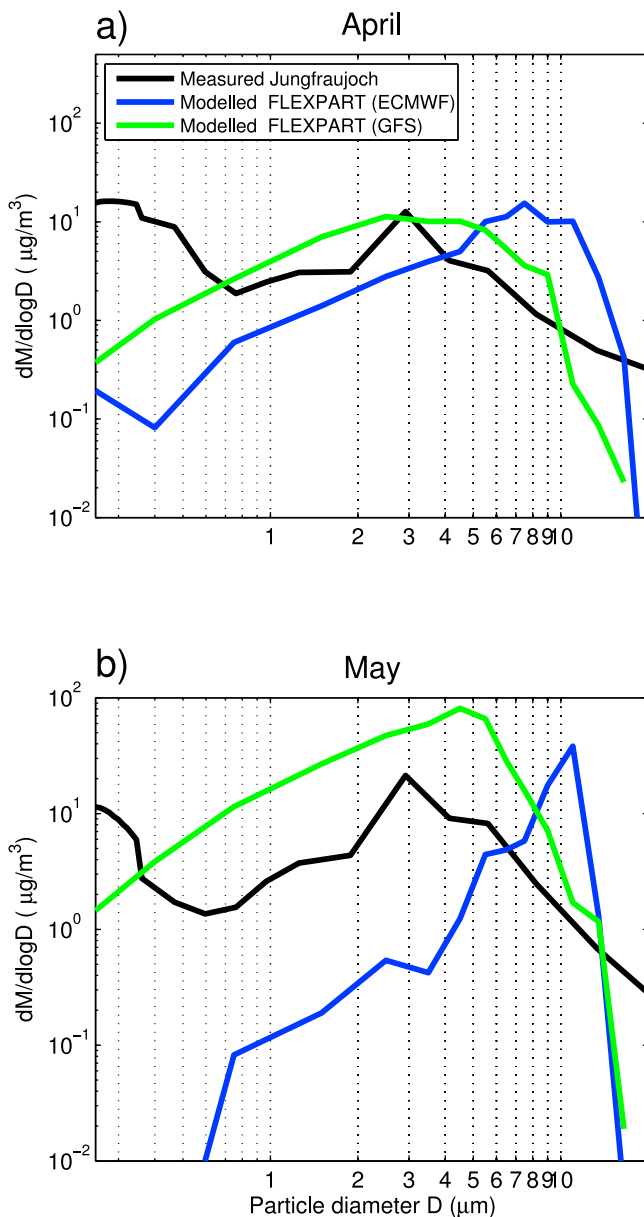
[74] The measured particle size distribution at Jungfraujoch peaks at  $\sim 3 \mu\text{m}$  (Figure 15) similar to the observations by the DIMO (Figure 13c) and the BAe-146 (Figures 5, 9a, and 13a). *Bukowiecki et al.* [2011] show that the efficiency of the inlet for particles was also very good for the larger particle sizes and that the particle size distributions were in agreement with other observations. Note that the peak in the measured size distribution at small particle sizes (diameter  $< 0.5 \mu\text{m}$ ) is ascribed to secondary aerosols (e.g., sulfate) which the models do not simulate. The modeled size distributions are rather different between the two FLEXPART simulations (ECMWF- and GFS-based) and the fit to the measured size distribution is not particularly good. In general, the modeled size distributions seem to be shifted to larger particle sizes, i.e., the modeled size distributions peak at a larger particle size than the measured distribution. The difficulty for the models to accurately simulate the ash cloud over the Alps is thus also seen clearly in the modeled size distributions.

### 3.2.3. Statistical Model-Measurement Analysis

[75] The dispersion model's performances in simulating the long-range ash transport are further assessed quantitatively by comparing paired measured and modeled values statistically. All the statistical indices used here are defined mathematically in Appendix A, and are described in more detail by *Mosca et al.* [1998]. A filter is applied to the paired measurement-model values in which only values for when measured concentrations are above  $10 \mu\text{g}/\text{m}^3$  are considered for the statistics. This captures the confirmed ash encounters

and thus the screened ash signal in the data. It is noted that filtering only according to measured values may favor models that over predict the width of the ash cloud and the time-interval of the concentration peaks.

[76] First, the full in situ BAe-146 and the Falcon data sets are used individually to evaluate the different models, and also to compare the two measurement data sets. Table 2 shows statistics for the seven different model simulations (FLEXPART ECMWF, FLEXPART GFS and NAME run with both a priori and a posteriori source terms, and the FLEXPART ECMWF with a uniform source term) and the mean over the three a posteriori simulations and the three a priori simulations, compared to the two in situ observation data sets. The time-integrated concentrations over the whole data set and the maximum ash concentrations of each measurement data set show that the models, and in particular the a posteriori simulations, are mostly able to capture the ash signals seen in the measurement data. The a posteriori simulations clearly put more mass at flight altitude than the a priori simulations (despite the reduction of ash release rate compared to the a priori as seen in Figure 1), but as seen in the case studies previously presented, the peak values are underestimated. The biases also indicate a general under prediction of the models with respect to the measurements. The tendency for the models to underestimate or overestimate the concentrations are given by the Factor Of EXceedance (FOEX) which counts the number of events of over or under prediction. The a priori simulations have a



**Figure 15.** Measured and modeled ash particle size distributions for the Jungfraujoch measurement station (46.55°N, 7.99°E, 3580 m a.s.l.) averaged over (a) 18–20 April 2010 and (b) 18–19 May 2010. An ash particle density of 2.65 g/cm<sup>3</sup> is assumed for the volume to mass conversions of the measurements.

FOEX around −20 to −40% which means that more values are under predicted than are over predicted, while the a posteriori simulations have a FOEX closer to 0% suggesting fewer under predictions. There is an increased skill shown in all the statistical indices for the a posteriori simulations over the a priori simulations. The statistics for the simulation assuming a uniform vertical distribution of the ash in the eruption column show the lowest score, although the emissions were 3–4 times higher for this set-up.

[77] As an estimate of the bias between the different measurement data sets, we normalize the bias with the measured mean concentration of each observation data set

(“Bias/mean” in Table 2). The differences in these values for each data set and model simulation are not very large, which indicates that there are no large systematic differences between the measurement data sets. However, a tendency for the Falcon data being more negatively biased is noted. This could imply that the Falcon data are slightly biased high against the BAe-146 data.

[78] Furthermore, Table 3 shows statistical measures when using the full in situ observation data set combining all BAe-146, Falcon, DIMO and Jungfraujoch measurements for April and May 2010. The Figure of Merit in Time (FMT) evaluates the temporal trend of the time series and the overlap between the measured and model predicted concentrations. This index is influenced by peak values, time and duration of the ash cloud “events,” and is very sensitive to time shift between measured and modeled values. In the along-track time series for the aircraft, a time shift can also be considered as a cloud positioning error as the aircraft is moving and not at a fixed location. This index alone may be misleading and an accurate evaluation of the start time of the event, duration of nonzero concentrations, peak values and time of occurrence (as shown in Figures 3, 7, 11, and 14) must accompany the analysis. *Mosca et al.* [1998] give a model having a 50% FMT a good performance score. As seen in the previous sections there were several occasions when the modeled ash cloud was shifted in time and/or position to the observations. Thus, an FMT around 30–50% is considered as moderately good for this dispersion problem. The FMT increases for all a posteriori simulations over the a priori simulations and the uniform simulation has a lower score.

[79] The normalized mean square errors (NMSE) give information on the deviations, and if small there is a limited spread of the modeled concentrations around the measurements. The most striking difference between model simulations is seen in NMSE as it is very sensitive to differences between observed and measured values. NMSE is also affected by shifts in time and space between the modeled and measured concentrations, and differences in peak values have a stronger influence on NMSE than on other indices. The NMSEs for the a posteriori simulations for all models are substantially lower than for the a priori simulations and suggest that the models perform quite well in space and time. The biases indicate also for this full observation data set a general under prediction of the modeled concentrations with respect to the measurements.

[80] Pearson’s correlation coefficients (PCC), for the concentrations indicate that there is a moderate positive correlation between the a posteriori modeled and measured ash concentrations. Significant correlations of 0.36–0.48 are obtained for the a posteriori simulations except for the FLEXPART ECMWF (0.14). This is linked to the strong influence of the BAe-146 measurements on 14 May which is discussed in more detail later. The Spearman’s rank correlation coefficient (SRCC) is less sensitive than the PCC to strong outliers that are in the tails of both samples and does not depend on a linear relationship between the variables. The SRCCs are high (0.25–0.62) for all a posteriori simulations including the FLEXPART-ECMWF, indicating that the PCC is affected by peak concentration values (outliers).

[81] The correlations suggest moderately good model performance for the a posteriori simulations. The a priori correlations are lower and not significant. Figure 16 shows

**Table 2.** Statistical Measures for in Situ Observed and Modeled Ash Concentrations for All BAe-146 (70 Values) and Falcon Flights (29 Values) in April and May 2010<sup>a</sup>

	Model (Met Data)									
	FLEXPART (ECMWF)			FLEXPART (GFS)		NAME (MetUM)		Model Mean		Observed
Simulation	uniform	a priori	a posteriori	a priori	a posteriori	a priori	a posteriori	a priori	a posteriori	<b>BAe-146 Falcon</b>
Integrated concentration	<b>4200</b> <i>700</i>	<b>2200</b> <i>900</i>	<b>7200</b> <i>1700</i>	<b>4400</b> <i>1200</i>	<b>8500</b> <i>1300</i>	<b>2200</b> <i>800</i>	<b>8500</b> <i>1100</i>	<b>2900</b> <i>1000</i>	<b>8100</b> <i>1400</i>	<b>10500</b> <i>2900</i>
Peak concentration	<b>0</b> <i>43</i>	<b>5</b> <i>68</i>	<b>297</b> <i>182</i>	<b>4</b> <i>80</i>	<b>623</b> <i>96</i>	<b>116</b> <i>28</i>	<b>657</b> <i>74</i>	<b>42</b> <i>59</i>	<b>526</b> <i>117</i>	<b>1805</b> <i>290</i>
Bias	<b>-89</b> <i>-75</i>	<b>-118</b> <i>-69</i>	<b>-47</b> <i>-42</i>	<b>-87</b> <i>-57</i>	<b>-28</b> <i>-53</i>	<b>-118</b> <i>-71</i>	<b>-28</b> <i>-60</i>	<b>-108</b> <i>-66</i>	<b>-34</b> <i>-52</i>	
FOEX	<b>-29</b> <i>-33</i>	<b>-23</b> <i>-36</i>	<b>-11</b> <i>-22</i>	<b>-20</b> <i>-29</i>	<b>-7</b> <i>-19</i>	<b>-24</b> <i>-40</i>	<b>-1</b> <i>-29</i>	<b>-21</b> <i>-26</i>	<b>-3</b> <i>-16</i>	
Bias/mean	<b>-0.6</b> <i>-0.8</i>	<b>-0.8</b> <i>-0.7</i>	<b>-0.3</b> <i>-0.4</i>	<b>-0.6</b> <i>-0.6</i>	<b>-0.2</b> <i>-0.5</i>	<b>-0.8</b> <i>-0.7</i>	<b>-0.2</b> <i>-0.6</i>	<b>-0.7</b> <i>-0.6</i>	<b>-0.3</b> <i>-0.5</i>	

<sup>a</sup>Bold values are for the BAe-146 while italic values are for the Falcon. Integrated concentrations ( $\mu\text{g}/\text{m}^3$ ), measured peak concentrations ( $\mu\text{g}/\text{m}^3$ ) and modeled concentration at the time corresponding to the measured peak, the bias ( $\mu\text{g}/\text{m}^3$ ) between modeled and measured concentrations, Factor Of EXceedance (FOEX) in percent [Mosca *et al.*, 1998], bias normalized by the measured mean concentration are reported. The observations have an uncertainty of a factor 2.

a scatterplot of measured versus modeled values from the mean over the three a posteriori and a priori simulations, respectively. The PCCs are  $-0.02$  (a priori) and  $0.36$  (a posteriori), while the SRCCs are  $0.21$  (a priori) and  $0.55$  (a posteriori). The most striking difference is seen for those observed values between  $100$  and  $1000 \mu\text{g}/\text{m}^3$  for which the a priori simulations often have modeled values close to  $1 \mu\text{g}/\text{m}^3$ , these values are increased to  $10$ – $100 \mu\text{g}/\text{m}^3$  in the a posteriori simulations.

[82] The statistics are highly influenced by the two BAe-146 flights on 14 May which cover half of the values for the BAe-146 comparison data set. There were very high measured values on this flight which the models did not capture very well, and there seemed to be a temporal/positioning error for the modeled concentrations to the measured values (Figure 3). Both the high peak values and the time-shift in the time series influence the statistics. Also as indicated in section 3.2.1.1, there were uncertainties related to certain measurement-model comparisons for this flight (peak 2 in Figure 3) when compared to SEVIRI satellite data. When excluding the 14 May flights from the statistical analysis, the deviations decrease (NMSE is lowered to 2–3) for all the a

posteriori simulations. Also the correlation coefficients in general increase, in particular for the FLEXPART-ECMWF a posteriori simulation to a significant value (from  $0.14$  to  $0.64$ ), while the PCC for the FLEXPART-GFS a posteriori is reduced to  $0.19$  and not significant. The PCC of the NAME a posteriori is slightly reduced but still significant, indicating that NAME performed quite well for the 14 May case. Also the SRCC shows increased score except for NAME. The FMT increases to up to  $46\%$  for the FLEXPART-ECMWF. These results are more consistent with the previous judgment of the model's performances in the previous sections where the FLEXPART-GFS simulations seemed to be inferior to the other model simulations. This may be related in part to the lower resolution of the GFS meteorological data. Also Stohl *et al.* [2011] found no significant improvements for the FLEXPART-GFS a posteriori simulations over the a priori simulation when comparing to peak values measured by the Falcon.

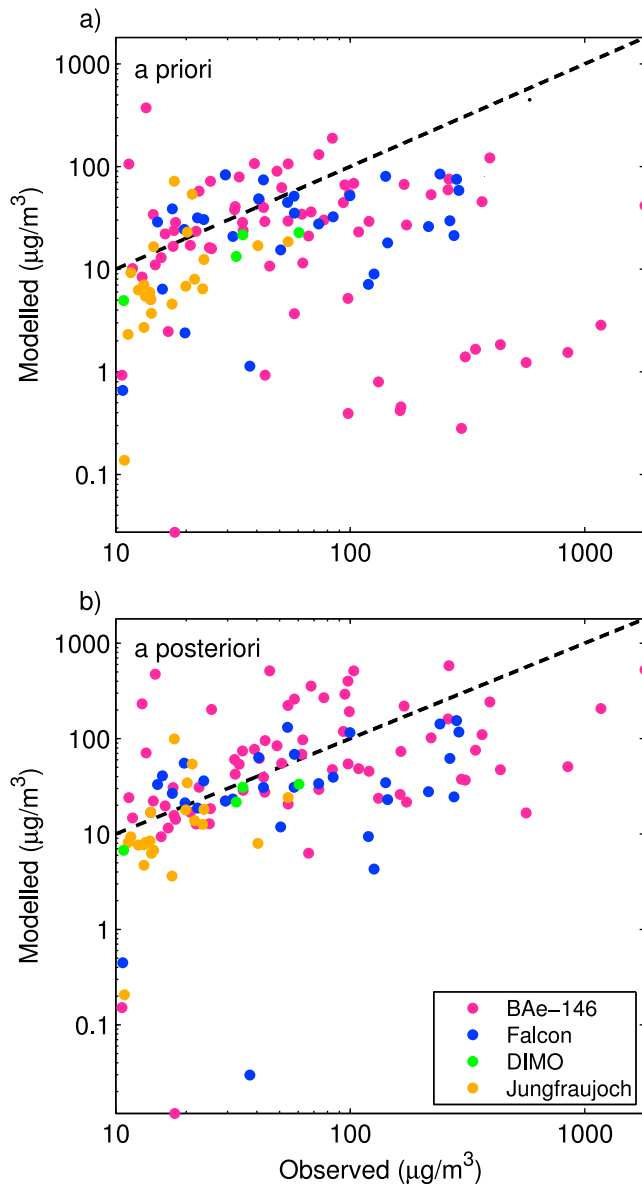
[83] Comparing and ranking the different models is challenging because the different statistical indices give different judgments. Also the previous sections showed different levels of diffusion for the models which means that the

**Table 3.** Statistical Comparison of 124 Paired in Situ Measured and Modeled Ash Concentrations for All BAe-146 (70 Values) and Falcon (29 Values) Flights in April and May, the DIMO Flight on 18 May (4 Values) and Measured PM10 Concentrations (21 Values) From Two Episodes in April and May at the Jungfraujoch Station<sup>a</sup>

	Model (Met Data)									
	FLEXPART (ECMWF)			FLEXPART (GFS)		NAME (MetUM)		Model Mean		
Simulation	uniform	a priori	a posteriori	a priori	a posteriori	a priori	a posteriori	a priori	a posteriori	
FMT	8 (16)	15 (30)	25 (46)	19 (27)	24 (30)	19 (31)	35 (39)	20 (35)	30 (43)	
NMSE	15 (8)	20 (3)	7 (2)	12 (5)	5 (3)	18 (4)	4 (3)	15 (3)	5 (2)	
Bias	-69 (2)	-85 (-24)	-38 (-20)	-62 (0)	-27 (-16)	-86 (-33)	-31 (-28)	-78 (-19)	-32 (-21)	
PCC	-0.08 (-0.04)	-0.08 (0.17)	0.14 (0.64)	-0.04 (0.04)	0.36 (0.19)	0.17 (0.14)	0.48 (0.39)	-0.02 (0.11)	0.36 (0.51)	
SRCC	-0.16 (0.22)	0.06 (0.37)	0.37 (0.48)	0.07 (0.34)	0.25 (0.39)	0.19 (0.44)	0.62 (0.51)	0.21 (0.48)	0.55 (0.53)	

<sup>a</sup>Figure of Merit in Time (FMT) [Mosca *et al.*, 1998], normalized mean square error (NMSE) and bias ( $\mu\text{g}/\text{m}^3$ ) between modeled and measured concentrations, Pearson's correlation coefficient (PCC) and Spearman's Rank correlation coefficient (SRCC) for the concentrations. Values in brackets are statistics when excluding the 14 May BAe-146 flight (leaving 90 paired values).





**Figure 16.** In situ measured ash concentrations from the BAe-146, Falcon and DIMO aircraft, and PM10 concentrations from the Jungfraujoch station versus modeled ash concentrations from the mean over the three simulations (FLEXPART-ECMWF, FLEXPART-GFS and NAME-MetUM) using the (a) a priori source term and (b) a posteriori source terms of Figure 1. The dashed black line shows the 1:1 line (perfect agreement). The Pearson's correlation coefficients (PCC) are  $-0.02$  (a priori) and  $0.36$  (a posteriori), while the Spearman's rank correlation coefficients (SRCC) are  $0.21$  (a priori) and  $0.55$  (a posteriori).

distribution of the modeled concentrations leading e.g., to better peak values, or to avoiding having a narrow plume in the wrong place mean that different models will do better on different statistics. Models that tend to be “smooth” (i.e., more dispersive) always outperform models with a “spiky” distribution in terms of FMT, NMSE, PCC, even though it is not clear they are “better.” It is very clear, however, that the a priori simulations and the simulation using a uniform source

term have a lower score than the a posteriori simulations. Given the uncertainties in the measurements (and other uncertainties), the performance of the models with these statistics, in particular the a posteriori simulations, is considered quite good.

#### 4. Discussions

[84] The source term estimated by the inversion (the a posteriori source term) showed that the fine ash was released mostly at the top of the eruption plumes as expected from simple models of buoyant plume rise. There exist no measurements of the actual source term to compare the a posteriori source term to. In the initial phase of the eruption (14 April) the emissions reached 10 km above the volcano but most of the emissions remained below 4 km. However, for the rest of the eruption, most of the significant ash emissions were above 3 to 4 km in altitude. When averaging over the three a posteriori source terms and considering the period 5–18 May when there were significant emissions, we find that the altitude of the maximum emission strength was on average at 5.3 km. Interestingly, *Arason et al.* [2011] report that the number of plume top altitudes estimated by the Keflavik radar peaks at 5–5.5 km altitude. Our a posteriori source terms for the 5–18 May period give that 54% of the ash emissions were released over the 2.5 km deep height range of the plume where the maximum emission strength occurred, whereas 8% were released above and 38% below this height range. Hence, methods for improving the uniform source profile commonly used operationally are clearly desirable. Here we have demonstrated such a method using an inversion technique.

[85] Even though the a posteriori source term in general showed lower emission rate of ash than the a priori source term, the a posteriori ash concentrations downwind of the volcano were mostly increased compared to the a priori simulations at the locations where high ash concentrations were observed. There are possibilities that biases in the satellite data might drive the a posteriori model concentrations from the inversion high, and improvements in the satellite retrievals are needed. However, the better agreement for the a posteriori concentrations with the measurements demonstrates that the a posteriori simulations released the ash in a more accurate time-window of the eruption period, and at a more accurate altitude than the a priori simulations. We also note that, the a priori description used is more sophisticated (based on PLUMERIA) than the common operational approach of using a uniform source profile. The FLEXPART model simulation assuming a uniform source mostly showed too much mass compared to vertical lidar profiles, while for the in situ measurement-comparisons the simulations on some occasions did not capture the observed ash signal (e.g., 14 May). This suggests that the ash emissions were not assigned to the right altitude and time-window of the eruption, and that the assumption of the fine ash fraction required for the uniform simulation set-up was uncertain. The assumption of the fine ash fraction (here 10%) varies significantly between eruptions [*Mastin et al.*, 2009] and is generally of high uncertainty. Satellite retrievals only include estimates of the fine ash mass left in the distal ash clouds and consequently this reduces the dependency on potentially erroneous a priori assumptions about the source and the

fraction of fine ash. Thus, using satellite retrievals in source term estimates as in the inversion method reduce the uncertainty related to the source term.

[86] The models generally underestimate the observed peak ash concentrations (in particular on the 14 May BAe-146 flights). This is most likely due to timing and positioning errors of the modeled ash cloud related to the modeled dispersion, and because the model output is averaged over time (one hour) and over a grid box which smooths out concentrations. Also, uncertainties and variations in the source term which are not accounted for in the 3 hourly resolution used here will affect the ability of the models to simulate peak concentrations. To capture the peak concentration values, introducing a buffer-zone to account for positional errors of the modeled cloud, and a peak-to-mean factor (accounting for the difference between modeled mean concentrations and peak concentrations) have been suggested and tested by Webster *et al.* [2012]. As the a posteriori ash concentrations showed closer agreement with the observed peak values, any adjustments to account for unresolved peak values will be less than for simulations using simplified source terms.

[87] The inversion using NAME model input gave a lower total emitted a posteriori mass than the two inversions based on FLEXPART model data (Table 1). Also the a posteriori simulations showed different total column values between the two models. Though there are many factors responsible for these differences, part of the discrepancies can be explained by differences in the two models regarding the removal of ash by wet deposition. Both NAME and FLEXPART distinguish between various precipitations situations (e.g., large-scale, convective, in-cloud (rain out) and below-cloud (wash out) scavenging). NAME also includes orographically enhanced precipitation and distinguishes between rain and snow. The two models use different wet scavenging coefficients for the calculations of wet deposition for the various precipitation events. Furthermore, the occurrence of clouds is calculated differently in the two models. Also, our simulations showed different levels of diffusion for the models which contributes to the model differences. Further evaluation of the model differences is needed in future studies.

[88] The initial particle size distribution used for the model simulations is a source of uncertainty related to the modeled ash transport. The real size distribution presumably varies for different volcanoes and eruptions, thus the initial size distribution at the source to be used for modeling is very difficult to estimate. For this study only a single initial size distribution was used for the simulations, however, the real size distribution will probably also vary during the eruption. Despite uncertainties, this study showed a reasonably good agreement between modeled and measured size distributions obtained from aircraft measurements. Phreatomagmatic eruptions like Eyjafjallajökull, where hot magma comes into contact with a large source of water, tend to generate more fine ash [Morrissey *et al.*, 2000] than other eruptions, and the fact that modeled particle size distributions showed too little mass in the 1–10  $\mu\text{m}$  particle size range compared to the measured distributions suggests that this was not adequately considered and reflects the uncertainty in the initial size distributions used for modeling. However, the role that the modeled particle size distribution has on the ash transport remains unclear as the modeled ash concentrations in

general compared reasonably well with the aircraft measured concentrations.

[89] Moreover, the fact that the models' particle size distributions for some model cases seemed to have too quick a drop-off for large particles might indicate that there are mechanisms that continuously aggregate particles, which the models do not simulate. Also the assumption of spherical particles in the model may yield too quick a drop-off for large particles. Varying shapes of the particles would mean that the too quick drop-off at large particle sizes is smoothed. Thus, a better characterization of the initial particle size distribution used for volcanic transport modeling, and how significant this parameter is for the ash transport, are important for future studies.

[90] The fact that the modeled particle size distributions seemed to be shifted to larger particle sizes, i.e., the models have more mass in the larger size classes than in the smaller ones, compared to the measured size distributions, may explain the shift in the vertical position of the ash cloud as seen in some of the lidar comparisons (Figures 4, 8 and 12). If the particles are too large in the model, the sedimentation may be overestimated. To examine the effect of too large particles we can calculate the fall distance a particle of 5  $\mu\text{m}$  diameter will have over a particle of size 10  $\mu\text{m}$ , within the time it takes for the particles to be transported from the volcano to the observation site. The Stoke's law [Hinds, 1999] relates the particle diameter  $D$ , the particle density  $\rho_p = 3.0 \text{ g/cm}^3$ , the gravitational acceleration  $g = 9.81 \text{ ms}^{-2}$  and the dynamic viscosity of air  $\mu = 14 \times 10^{-6} \text{ kgm}^{-1} \text{ s}^{-1}$ , to the particles' settling velocity (m/s):  $\nu = \frac{g\rho_p D^2}{18\mu}$ . The vertical distance an ash particle would fall from the time it is released from the volcano to when it reaches the observation site is found by multiplying the fall velocity with the age of the particle. Schumann *et al.* [2011] report that the age of the ash cloud observed by the Falcon on 17 May was approximately 72 h. A spherical particle with a diameter of 5  $\mu\text{m}$  will in 72 h fall approximately 750 m, while a larger particle with diameter of 10  $\mu\text{m}$  will fall about 3000 m in the same time. Also the effect of a too large particle density may contribute to this, as larger density causes a faster sedimentation. However, a FLEXPART simulation with reduced particle density from 3.0  $\text{g/cm}^3$  to 2.3  $\text{g/m}^3$  did not place the ash clouds at lower altitudes (not shown). A further complicating factor is the shape of the particles, which are not spherical but angular. Particles with varying shapes tend to take longer to fall. Thus, a size distribution shifted to large particles, and the assumption of spherical particles in the model may place the modeled ash cloud at a too low altitude compared to the observations.

[91] The models clearly have limitations which are important to be aware of. The uncertainty in the model predictions increases as the ash cloud is transported further away from its source. Furthermore, the transport of an inhomogeneous ash cloud over complex topography is especially difficult to model accurately. The modeled ash clouds were shifted in time compared to surface measurements at Jungfraujoch, indicating too slow transport to the station. Also the comparison to aircraft measurements over the Alps showed less agreement than to other aircraft observations closer to the volcano. These limitations are important to keep in mind when evaluating volcanic ash forecasts.

[92] The a posteriori model forecast improvements seen in this study are due in part to constraints supplied by SEVIRI mass loading retrievals. In northern Europe, including Iceland, SEVIRI data are available every 5 min thus offering a wonderful opportunity to utilize SEVIRI data operationally. Including data as close in time as possible to the model initialization time is sensible. For example, a posteriori results for 14 May might be improved by using only data from 15 May. However, in a forecast situation we can only utilize satellite data up until the current time, but as more satellite data becomes available the inversion can be re-run to give a re-estimated and improved the source term.

## 5. Conclusions

[93] In this paper, a detailed assessment of dispersion model performance for volcanic ash transport has been given for the eruption of the Eyjafjallajökull volcano in April and May 2010. Two different dispersion models, FLEXPART and NAME, run on different meteorological data for driving the simulations (ECMWF, GFS and MetUM) have been tested for modeling the ash concentrations during the eruption. The following are the main findings from this study:

[94] 1. Source terms for the ash emissions as a function of height and time were estimated with an inversion algorithm constraining a priori emissions by satellite data (SEVIRI). The a posteriori source term from the inversions gave a mean total emission of  $7.6 \pm 4.8$  Tg of fine ash ( $2.8\text{--}28\ \mu\text{m}$ ) released into the atmosphere during the whole eruption period, with two main periods of strong ash emissions (14–16 April and 5–18 May). The a posteriori source terms differed significantly from the a priori source term in that the emissions were more defined as strong pulses, releasing the ash mostly near the top of the eruption plumes. In May, 54% of the ash emissions were released over the 2.5 km deep height range of the plume where the maximum emission strength occurred, whereas 8% were released above and 38% below this height range.

[95] 2. The similarity of the source terms derived from inversions using input from different dispersion models, run on various meteorological data, demonstrated that the source terms were robust to which dispersion model and meteorological data were used.

[96] 3. Long-range transport simulations of the ash emissions were compared to a large set of observations, both surface and airborne. Measurements from three research aircraft (FAAM BAe-146, DLR Falcon and the Swiss MetAir DIMO), together with measurements from the measurement station Jungfraujoch, showed that the modeled ash concentrations using the a posteriori source term from the inversion corresponded well with observations and was mostly within the factor two uncertainty of the aircraft measurements. The vertical positioning of the a posteriori modeled ash clouds agreed to  $\pm 1$  km with lidar profiles of the ash clouds taken from the research aircraft. Also, despite uncertainties, this study showed a reasonably good agreement between modeled and measured size distributions obtained from aircraft measurements.

[97] 4. On some occasions, there were large differences in both positioning and amount of ash at certain locations between the different model simulations initiated with the a posteriori source terms, despite the similarity of the source terms used. In particular, large uncertainties in the model simulations were found for transport of an inhomogeneous

ash cloud over complex topography (e.g., the Alps). An ensemble of models and input meteorology could be beneficial in predicting ash cloud movements and ash concentrations with improved quantification of uncertainties.

[98] 5. The a posteriori simulations showed better skill when compared statistically to the observations than the simulations using the a priori source term or simpler source term with uniform vertical distribution. Correlation between observed and measured ash concentrations increased from  $-0.02$  (a priori) and  $-0.08$  (uniform) to a significant moderate positive correlation  $0.36$  (a posteriori). Also the deviations (NMSE) were clearly reduced suggesting that the models perform reasonably well in space and time. The Figure of Merit in Time (FMT) score increased to up to 46% indicating a good model performance. The simulation assuming a uniform distribution of the ash within the eruption column yielded less accurate results.

[99] 6. Overall, the improvements in the a posteriori simulations over the a priori simulations demonstrate that the inversion provides invaluable information that, on a near-real time basis, can be used as input to the dispersion models without the need for human intervention, and that this will improve quantitative ash forecasts. Further studies are planned for other volcanoes, in other parts of the world, which will allow a more detailed assessment of the prediction improvements.

## Appendix A

[100] This appendix defines the statistical terms used in this paper and describes their meaning briefly. A more detailed review of these statistical terms is given by Mosca *et al.* [1998].

[101] The bias is defined as the average difference between paired modeled predicted,  $P_i$ , and measured,  $M_i$ , values:

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (P_i - M_i), \text{ where } N \text{ is the number of pairs } (M_i, P_i).$$

The bias is an estimation of the general over prediction or under prediction of the model with respect to the measurements.

[102] The Normalized Mean Square Error is defined as

$$\text{NMSE} = \frac{1}{N} \sum_{i=1}^N (P_i - M_i)^2 / \overline{PM}, \text{ where } \overline{P} = \frac{1}{N} \sum_{i=1}^N P_i \text{ and}$$

$$\overline{M} = \frac{1}{N} \sum_{i=1}^N M_i \text{ are the average model predictions and mea-}$$

surements, respectively. The NMSE gives information on the deviations. A model with a very low NMSE is performing well both in space and time.

[103] The Factor Of EXcedance is defined as  $\text{FOEX} = \left[ \frac{N_{(P_i > M_i)}}{N} - 0.5 \right] \times 100$ , where  $N_{(P_i > M_i)}$  is the number of pairs where the model predicted value is greater than the measured value, i.e., the number of over predictions. FOEX ranges between  $-50$  (all model values are under predicted) and  $+50\%$  (all model values are over predicted), and indicates whether over predictions or under predictions are more frequent.

[104] The Figure of Merit in Time is defined as  $\text{FMT} =$

$$100 \times \frac{\sum_{j=1}^N \min\{M(\bar{x}, t_j), P(\bar{x}, t_j)\}}{\sum_{j=1}^N \max\{M(\bar{x}, t_j), P(\bar{x}, t_j)\}}, \text{ where } M(\bar{x}, t_j) \text{ represents the}$$

measured concentration at the same location,  $\bar{x}$ , and at the same time,  $t_j$ , as the model predicted concentration  $P(\bar{x}, t_j)$ . The FMT evaluates the temporal trend of the overlap between the measured and model predicted concentrations of the time series. The FMT is normalized to the maximum predicted or measured value at each time interval and is expressed as a percentage value.

[105] The Pearson's correlation coefficient is also called the linear correlation coefficient and is defined as

$$PCC = \frac{\sum_{i=1}^N (M_i - \bar{M})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^N (M_i - \bar{M})^2} \sqrt{\sum_{i=1}^N (P_i - \bar{P})^2}}. \text{ A model with } PCC = 111$$

has a complete correlation between model predicted and measured values. The Spearman's rank correlation coefficient (SRCC) has the same mathematical definition as the PCC, but calculated between the *ranked* variables. The values are ranked according to their position in the ascending order of the values.

[106] **Acknowledgments.** This work was funded by the European Space Agency's Earth Observation Envelope Programme (EOEP) – Data User Element through the project Support for Aviation for Volcanic Ash Avoidance (SAVAA). Airborne data was obtained using the BAe-146-301 Atmospheric Research Aircraft (ARA) flown by Direct Flight Ltd and managed by the Facility for Airborne Atmospheric Measurements (FAAM), which a joint entity of the Natural Environment Research Council (NERC) and the Met Office. The staff of the Met Office, FAAM, Direct Flight, Avalon Engineering and BAE Systems are thanked for their dedication in making the measurement campaign a success. The Falcon flights were performed on request of the Deutscher Wetterdienst (DWD, German Weather Service) and the Bundesministerium für Verkehr, Bau und Stadtentwicklung (BMVBS, Federal Ministry of Transport, Building and Urban Development). We thank the International Foundation High Altitude Research Stations Jungfraujoch and Gomsgrat (HFSJG) for the opportunity to perform experiments on the Jungfraujoch.

## References

- Alaska Volcano Observatory (2010), Puff-Volcanic Ash Tracking Model, Eyjafjöll Volcano, Sept. 2011, [http://puff.images.alaska.edu/watch\\_Eyjafjoll.shtml](http://puff.images.alaska.edu/watch_Eyjafjoll.shtml), Anchorage, Alaska.
- Arason, P., G. N. Petersen, and H. Björnsson (2011), Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010, *Earth Syst. Sci. Data*, 3, 9–17, doi:10.5194/essd-3-9-2011.
- Barcelona Supercomputing Center (2010), Eyjafjalla Fall3d simulations, Sept. 2011, <http://www.bsc.es/projects/earthscience/nube2010/>, Barcelona, Spain.
- Bukowiecki, N., et al. (2011), Ground-based and airborne in-situ measurements of the Eyjafjallajökull volcanic aerosol plume in Switzerland in spring 2010, *Atmos. Chem. Phys.*, 11, 10,011–10,030, doi:10.5194/acp-11-10011-2011.
- Costa, A., A. Folch, and G. Macedonio (2010), A model for wet aggregation of ash particles in volcanic plumes and clouds: 1. Theoretical formulation, *J. Geophys. Res.*, 115, B09201, doi:10.1029/2009JB007175.
- Dacre, H. F., et al. (2011), Evaluating the structure and magnitude of the ash plume during the initial phase of the 2010 Eyjafjallajökull eruption using lidar observations and NAME simulations, *J. Geophys. Res.*, 116, D00U03, doi:10.1029/2011JD015608.
- Devenish, B. J., P. N. Francis, B. T. Johnson, R. S. J. Sparks, and D. J. Thomson (2012), Sensitivity analysis of dispersion modeling of volcanic ash from Eyjafjallajökull in May 2010, *J. Geophys. Res.*, doi:10.1029/2011JD016782, in press.
- De Wekker, S. F. J., D. G. Steyn, and S. Nyeki (2004), A comparison of aerosol layer- and convective boundary layer structure over a mountain range during STAAARTE '97, *Boundary Layer Meteorol.*, 113, 249–271, doi:10.1023/B:BOUN.0000039371.41823.37.
- Eckhardt, S., A. J. Prata, P. Seibert, K. Stebel, and A. Stohl (2008), Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling, *Atmos. Chem. Phys.*, 8, 3881–3897, doi:10.5194/acp-8-3881-2008.
- European Commission (2010), Volcano crisis report, online report, Brussels. [Available at [http://ec.europa.eu/transport/ash\\_cloud\\_crisis\\_en.htm](http://ec.europa.eu/transport/ash_cloud_crisis_en.htm).]
- European Commission (2011), Volcano Grimsvötn: How is the European response different to the Eyjafjallajökull eruption last year? Frequently Asked Questions, 26 May 2011, *MEMO/11/346*, Brussels. [Available at [http://ec.europa.eu/transport/ash\\_cloud\\_crisis\\_en.htm](http://ec.europa.eu/transport/ash_cloud_crisis_en.htm).]
- Folch, A., A. Cosat, A. Durant, and G. Macedonio (2010), A model of wet aggregation of ash particles in volcanic plumes and clouds: 2. Model application, *J. Geophys. Res.*, 115, B09202, doi:10.1029/2009JB007176.
- Hinds, W. C. (1999), *Aerosol Technology*, 504 pp., John Wiley, New York.
- International Civil Aviation Organization (2007), Manual on volcanic ash, radioactive material, and toxic chemical clouds, 2nd ed., *Doc 9691-AN/954*, Montreal, Que., Canada. [Available at <http://www.paris.icao.int/news/pdf/9691.pdf>.]
- Jakobsdóttir, S. S., G. Sigursson, and S. Hreinsdóttir (2010), Eruption in Eyjafjallajökull, status reports, Sept. 2011, [http://www.earthice.hi.is/page/ies/Eyjafjallajokull\\_eruption?224.29](http://www.earthice.hi.is/page/ies/Eyjafjallajokull_eruption?224.29), Univ. Icel., Reykjavík.
- Jones, A., D. Thomson, M. Hort, and B. Devenish (2007), The UK Met Office's next generation atmospheric dispersion model, NAME III, in *Air Pollution Modelling and its Application XVII*, edited by C. Borrego and A.-L. Norman, pp. 580–589, Springer, New York.
- Kristiansen, N. I., et al. (2010), Remote sensing and inverse transport modeling of the Kasatochi eruption sulfur dioxide cloud, *J. Geophys. Res.*, 115, D00L16, doi:10.1029/2009JD013286.
- Leadbetter, S., and M. Hort (2011), Volcanic ash hazard climatology for an eruption of Hekla Volcano, Iceland, *J. Volcanol. Geotherm. Res.*, 199, 230–241, doi:10.1016/j.jvolgeores.2010.11.016.
- Marenco, F., B. Johnson, K. Turnbull, S. Newman, J. Haywood, H. Webster, and H. Ricketts (2011), Airborne lidar observations of the 2010 Eyjafjallajökull volcanic ash plume, *J. Geophys. Res.*, 116, D00U05, doi:10.1029/2011JD016396.
- Maryon, R. H., D. B. Ryall, and A. L. Malcolm (1999), The NAME 4 dispersion model: Science documentation, *Met Off. Turbul. and Diffus. Note 262*, Natl. Meteorol. Libr., Exeter, U. K.
- Mastin, L. G. (2007), A user-friendly one-dimensional model for wet volcanic plumes, *Geochem. Geophys. Geosyst.*, 8, Q03014, doi:10.1029/2006GC001455.
- Mastin, L. G., et al. (2009), A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions, *J. Volcanol. Geotherm. Res.*, 186, 10–21, doi:10.1016/j.jvolgeores.2009.10.013.
- Monitoring Atmospheric Composition and Climate (2010), Eruption of the Eyjafjallajökull volcano, Sept. 2011, [http://www.gmes-atmosphere.eu/news/volcanic\\_ash/background/](http://www.gmes-atmosphere.eu/news/volcanic_ash/background/), European Union, Brussels.
- Monks, P. S., et al. (2009), Atmospheric composition change–Global and regional air quality, *Atmos. Environ.*, 43, 5268–5350, doi:10.1016/j.atmosenv.2009.08.021.
- Morrissey, M., B. Zimanowski, K. Wohletz, and R. Buettner (2000), Phreatomagmatic fragmentation, in *Encyclopedia of Volcanoes*, edited by H. Sigurdsson, pp. 431–445, Academic, San Diego, Calif.
- Mosca, S., G. Graziani, W. Klug, R. Bellasio, and R. Bianconi (1998), A statistical methodology for the evaluation of long-range dispersion models: An application to the ETEX exercise, *Atmos. Environ.*, 32(24), 4307–4324, doi:10.1016/S1352-2310(98)00179-4.
- Näslund, E., and L. Thaning (1991), On the settling velocity in a nonstationary atmosphere, *Aerosol Sci. Technol.*, 14, 247–256, doi:10.1080/02786829108959487.
- National Environmental Research Institute (2010), Dispersion of the volcanic plume, Sept. 2011, <http://www.dmu.dk/International/Air/Models/volcanicplume/>, Roskilde, Denmark.
- Newhall, C. G., and S. Self (1982), The Volcanic Explosivity Index (VEI) An estimate of explosive magnitude for historical volcanism, *J. Geophys. Res.*, 87, 1231–1238, doi:10.1029/JC087iC02p01231.
- NILU - Norwegian Institute for Air Research (2010), Special Forecast products for Eyjafjallajökull produced through the SAVAA project, Sept. 2011, <http://transport.nilu.no/products/eyjafjallajokull>, Kjeller, Norway.
- Norwegian Meteorological Institute (2010), SNAP volcanic ash transport simulations, Sept. 2011, <http://retro.met.no/eyjafjalla/english.shtml>, Oslo.
- Prata, A. J., and A. Tupper (2009), Aviation hazards from volcanoes: The state of the science, *Nat. Hazards*, 51(2), 239–244, doi:10.1007/s11069-009-9415-y.
- Prata, A. J., S. A. Carn, A. Stohl, and J. Kerkmann (2007), Long range transport and fate of a stratospheric volcanic cloud from Soufrière Hills volcano, Montserrat, *Atmos. Chem. Phys.*, 7(19), 5093–5103, doi:10.5194/acp-7-5093-2007.
- Schumann, U., et al. (2011), Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010, *Atmos. Chem. Phys.*, 11, 2245–2279, doi:10.5194/acp-11-2245-2011.

- Seibert, P. (2000), Inverse modelling of sulfur emissions in Europe based on trajectories, in *Inverse Methods in Global Biogeochemical Cycles*, *Geophys. Monogr. Ser.*, vol. 114, edited by P. Kasibhatla et al., pp. 147–154, AGU, Washington, D. C., doi:10.1029/GM114p0147.
- Seibert, P., N. I. Kristiansen, A. Richter, S. Eckhardt, A. J. Prata, and A. Stohl (2011), Uncertainties in the inverse modelling of sulphur dioxide eruption profiles, *Geomatics Natl. Hazards Risk*, 2(3), 201–216, doi:10.1080/19475705.2011.590533.
- Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. S. Gilbert, L. S. Glaze, H. Sigurdsson, and A. W. Woods (1997), *Volcanic Plumes*, 574 pp., John Wiley, Chichester, N. Y.
- Stohl, A., M. Hittenberger, and G. Wotawa (1998), Validation of the Lagrangian Particle Dispersion model FLEXPART against large-scale tracer experiment data, *Atmos. Environ.*, 32(24), 4245–4264, doi:10.1016/S1352-2310(98)00184-8.
- Stohl, A., C. Forster, A. Frank, P. Seibert, and G. Wotawa (2005), Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5(9), 2461–2474, doi:10.5194/acp-5-2461-2005.
- Stohl, A., et al. (2009), An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, *Atmos. Chem. Phys.*, 9, 1597–1620, doi:10.5194/acp-9-1597-2009.
- Stohl, A., et al. (2011), Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: The 2010 Eyjafjallajökull eruption, *Atmos. Chem. Phys.*, 11, 4333–4351, doi:10.5194/acp-11-4333-2011.
- Turnbull, K. F., B. T. Johnson, F. Marengo, J. M. Haywood, A. Minikin, B. Weinzierl, H. Schlager, U. Schumann, S. J. Leadbetter, and A. M. Woolley (2012), A case study of observations of volcanic ash from the Eyjafjallajökull eruption, Part I: In situ airborne observations, *J. Geophys. Res.*, doi:10.1029/2011JD016688, in press.
- Volcanic Ash Advisory Centre London (2010), Volcanic Ash Advisory from London—Issued graphics, Sept. 2011, [http://www.metoffice.gov.uk/aviation/vaac/vaacuk\\_vag.html](http://www.metoffice.gov.uk/aviation/vaac/vaacuk_vag.html), Exeter, U. K.
- Webster, H. N., and D. J. Thomson (2012), Dry deposition modelling in a Lagrangian dispersion model, *Int. J. Environ. Pollut.*, in press.
- Webster, H. N., et al. (2012), Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, 117, D00U08, doi:10.1029/2011JD016790.
- Weigel, A. P., F. K. Chow, and M. W. Rotach (2007), The effect of mountainous topography on moisture exchange between the “surface” and the free atmosphere, *Boundary Layer Meteorol.*, 125(2), 227–244, doi:10.1007/s10546-006-9120-2.
- Wen, S., and W. I. Rose (1994), Retrieval of sizes and total masses of particles in volcanic clouds using AVHRR bands 4 and 5, *J. Geophys. Res.*, 99, 5421–5431, doi:10.1029/93JD03340.
- Witham, C. S., M. C. Hort, R. Potts, R. Servranckx, P. Husson, and F. Bonnardot (2007), Comparison of VAAC atmospheric dispersion models using the 1 November 2004 Grimsvötn eruption, *Meteorol. Appl.*, 14, 27–38, doi:10.1002/met.3.
- N. Bukowiecki, Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, CH-5232 Villigen, Switzerland.
- H. Dacre, Department of Meteorology, University of Reading, Reading RG6 6BB, UK.
- S. Eckhardt, N. I. Kristiansen, A. J. Prata, and A. Stohl, NILU - Norwegian Institute for Air Research, PO Box 100, N-2027 Kjeller, Norway. (nik@nilu.no)
- S. Henne, Empa, Swiss Federal Laboratories for Materials Science and Technology, CH-8600 Dübendorf, Switzerland.
- M. C. Hort, B. T. Johnson, F. Marengo, D. J. Thomson, and H. N. Webster, Met Office, FitzRoy Road, Exeter EX1 3PB, UK.
- B. Neining, Centre for Aviation, Zurich University of Applied Sciences, CH-8401 Winterthur, Switzerland.
- O. Reitebuch and B. Weinzierl, Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, D-82234 Oberpfaffenhofen, Germany.
- P. Seibert, Institute of Meteorology, University of Natural Resources and Life Sciences, A-1190, Vienna, Austria.